QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLGY

VOL. XII
1960

THE

QUARTERLY JOURNAL
OF
EXPERIMENTAL PSYCHOLOGY

Volume XII

PUBLISHED FOR
THE EXPERIMENTAL PSYCHOLOGY SOCIETY
BY
W. HEFFER & SONS LTD. CAMBRIDGE

The Quarterly Journal of Experimental Psychology

VOLUME XII, 1960

CONTENTS

SHORTER ARTICLES AND NOTES

APPARATUS

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XII 1960 Part 1

THE INFORMATION-CAPACITY OF THE HUMAN MOTOR-SYSTEM IN PURSUIT TRACKING

BY

E. R. F. W. CROSSMAN

From the Department of Psychology, Reading University

Human performance at a typical pursuit tracking task has been analysed by information theory. Without preview of the course, the channel-capacity was found to be about 4 bits per sec. With preview performance ceased to be perfect at about 2 bits per sec. but continued to improve up to at least 8 bits per sec., at about half the ideal slope. The findings are discussed in relation to the information-capacities measured in other tasks, and it is suggested that with full preview the rate should reach a maximum of about 10 bits per sec. Without preview the rate appears to be lowered by the need to offset response lag by predicting course behaviour.

Introduction

The usual experimental pursuit-tracking task consists of a display with a pointer which moves from side to side providing a course or input, $i(t)$. The subject has a control connected to a controlled member, a pointer or pen, which produces a track or output, $o(t)$, and his task is to keep the controlled member aligned with the course pointer so that the error, $e(t)$, is a minimum where

$$e(t) = o(t) - i(t) \qquad \text{(1)}$$

In engineering terms the subject is then part of a closed-loop control system or servo-mechanism whose function is to "follow-up" the course-pointer. There is a well-developed mathematical theory of linear servomechanisms and one of its central concepts is the *transfer-function*, which gives $o(t)$ as a function of $e(t)$ and its differential coefficients; hence several workers have attempted to specify the human transfer-function and they have had some success (Hick and Bates, 1951; North, 1950). Others (e.g. Fitts, Noble and Warren, 1955) thinking along similar lines have measured subjects' "frequency-response" and "phase-lag" when following sinusoidal inputs. However, this approach does not take account of the fact that the human subject has the ability, not shared by engineering control devices, to learn the pattern of the course and hence predict ahead, a fact which has been demonstrated by Poulton (1952, 1957) and called by him "perceptual anticipation." Nor does it allow for his ability to behave at will as an open-chain controller, by attending to the course itself rather than to the error, or to select certain parts of the course for special attention.

The latter aspects of performance can only be taken into account by considering the *statistical* rather than the functional properties of input, output and error, a step suggested by North (1950) but not followed up experimentally. Information-theory is the natural tool to choose for this purpose, and one is led to ask, not what is the output as a function of the error, but at what rate is information transmitted from

input to output, as a function of the input entropy. The important quantities to study are not then the actual functions $i(t)$, $o(t)$ and $e(t)$, but the time-averaged distributions of their values, $p(i)$, $p(o)$ and $p(i,o)$, and the entropies $H(i)$, $H(o)$ and $H(i,o)$ which measure their predictability. The subject's performance is measured by the rate R at which he can accept information from the course and reproduce it as track where—

$$R = H(i) + H(o) - H(i,o) \qquad \text{(2)}$$

The quantities in equation (2), being parameters of a continuous signal or ensemble of signals, must be handled according to the theory of continuous information, which differs in important respects from the discrete theory hitherto chiefly used in psychology.

The Theory of Continuous Information

(a) The dependence of information-rate on the choice of units.

In a discrete information-system the entropy of a signal depends only on the relative frequencies with which its different states occur, but in a continuous system it depends also on the size of the step, unit, grouping interval, or quantum used to measure it, increasing with the smallness of this unit (see Shannon and Weaver, 1949, Chap. III). However, provided that all the entropies are computed in the same terms and that the unit is small compared to the finest detail in the probability-distribution of the variable being considered, the rate is independent of it.

If the unit or grouping interval is too large, genuine fine-structure may be blurred out and the rate will be underestimated. But if it is too small, a large sample of performance is needed to give a good estimate of the probability-distribution. If the distribution is estimated from a small sample, chance differences of frequency due to sampling error appear as spurious fine detail and the rate is overestimated. The size of this bias has been calculated (Draper, 1954) and a correction can be applied, but the choice of unit must still be a compromise between the two conflicting requirements. For a given size of unit the rate has a definite maximum, equal to the input entropy.

(b) Time-sampling

The theory of continuous information sets a limit to the fineness with which the time-scale need be divided up. Since variables in the physical world cannot change instantaneously, their successive values at short time-intervals are highly correlated, which reduces the average information per value; and it has been shown (Shannon and Weaver, 1948, p. 53) that all the information contained in a given variable can be extracted by sampling it at least twice per period of the highest frequency W it can contain (i.e. its bandwidth), that is at intervals Δ, where

$$\Delta = \frac{1}{2W} \qquad \text{(3)}$$

(c) Distortion and noise

A distinction is drawn in the theory between two possible sources of discrepancy between the input and output signals in a channel, namely *distortion* and *noise*, and only the latter is considered as a loss of information. Distortion is any discrepancy due to a known systematic effect, and it can in principle be corrected at the receiving end without loss of information; the remainder, due to unpredictable effects, is noise which does cause loss of information. Distortion arose in the present experiments from delay in the subject's response and from a loss of amplitude at high speeds, and both were corrected before calculating the rate.

These and other theoretical considerations indicate that a tracking experiment must fulfil certain conditions if the calculated information-rate is to be meaningful. First, the course and the subject's performance must both be statistically "stationary" during a trial, so that the averaging processes shall be valid. Second, the successive samples of the course must be, or appear to the subject to be, uncorrelated, so that the calculated input entropy shall fairly represent what he experiences. Third, the bandwidths of course and track must be known and similar, so that the proper sampling interval for both can be

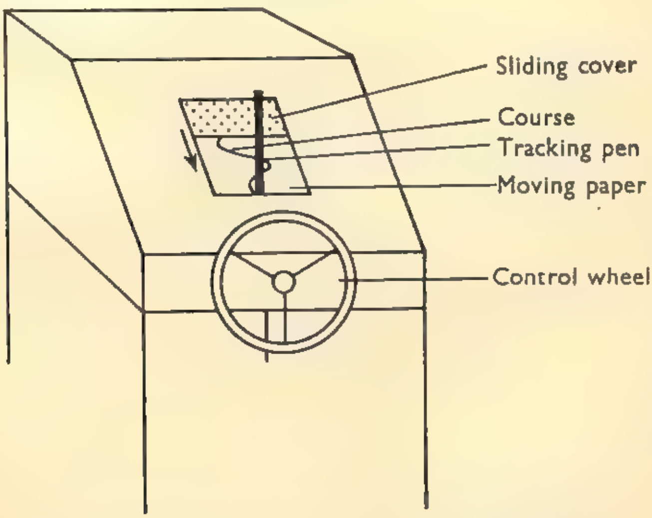
calculated from equation (3). Having obtained a record of the subject's performance the following steps must then be taken to obtain his information-rate:—

(1) Measure the values of corresponding sample-points of course and track at time-intervals Δ chosen according to the bandwidth.
(2) Compute the frequency-distributions of the sample values of course and track and their joint distribution, with a properly chosen grouping interval, allowing if necessary for distortion.
(3) Compute the input, output and joint entropies per sample-point.
(4) Compute the rate per sample from equation (2), and hence the rate per sec.

Experimental Method

The apparatus used for the present experiments is illustrated in Figure 1. The main unit consisted of a paper-drive and course-generator driven by a variable speed velodyne motor, and the control was an 18 in. vertical hand-wheel geared directly to a laterally moving pen, the controlled member. The subject stood facing a window within which he could see the course drawn by a hidden pen in black ink on a band of paper 8·5 cm. wide and his own pen made a red line. The amount of the course visible above his pen, the *preview*, could be reduced by a sliding blind; both course and track were visible for a short distance below the subject's pen. The hand-wheel had to be moved through about 40° to cover the full width of the track and the subject's hand-movements thus extended about 6 in. up or down in a slight curve; the wheel offered little resistance. For a trial the motor was started by a switch after a verbal warning, and the paper then ran at set speed until one cycle of the course-generator had been completed, when it stopped automatically. Eight speeds were used; they are given in Table III.

FIGURE 1



The apparatus used in the tracking experiment.

The subjects were instructed to do their best to keep the pen-point on the course, but that if unable to follow all its variations they were to attempt to reproduce it even if late. Four subjects were tested in the main experiment; each was given two practice runs, then one experimental run at each of the eight speeds with short (½ cm.) and long (8 cm.) preview, making 16 runs in all, the conditions being taken in random order. For the first two subjects the course was a pair of black lines 1 cm. apart, and for the second two a single line. The time-variation of the course was generated by a pair of epicyclic gears driven from the same shaft as the paper-drive, so that the course had the same shape on the paper at any speed. The gears had ratios of 5/3 and 3/4 and slightly different radii, the course function being:—

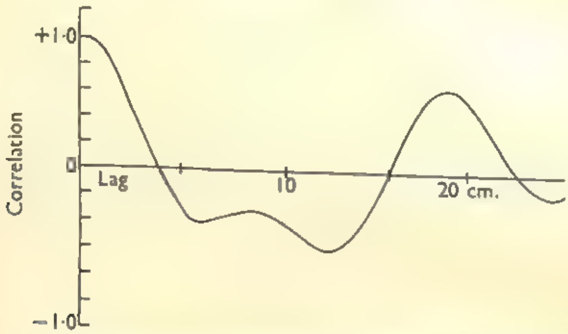
$$i(x) = 1{\cdot}25 \sin x + 1{\cdot}10 \sin \frac{8}{3} x + 1{\cdot}27 \sin \frac{17}{12} x$$

where x = distance along the paper in units of 3·8 cm.

The h.c.f. of the three frequency-components in the course is $\frac{1}{12}$ and it therefore repeated after respectively 12, 32 and 17 cycles. The apparatus automatically stopped after one such repetition, which constituted a run. There were 94 peaks (reversals of direction) during one run.

The course was intended to be an approximation to a band-limited random signal, in that the subjects were not supposed to be able to predict ahead by learning the pattern. Its autocorrelation-function, which is very similar to a part of the course itself, is given in Figure 2, and it can be seen that the maximum correlation in the first 25 cm. is $r = +0{\cdot}64$; r does not reach unity until at the end of one complete cycle, so that subjects cannot in principle improve their performance greatly except by learning the course as a whole, which did not appear to happen. It was therefore considered reasonable to regard the course as unpredictable over any range greater than 2 cm.

FIGURE 2



The autocorrelation function for the course used.

Analysis of Performance

The bandwidth of the course was taken to equal the highest frequency present in it, whose wavelength was 8·8 cm. on the paper; by the sampling theorem, equation (3), it would therefore be sufficient to sample its value every 4·4 cm. Since the subject's finite response-time always caused the track to lag behind the course, and it was desired to correct this distortion by comparing them at equivalent rather than simultaneous points, it was necessary to have readily recognizable sample-points. The successive peaks on the course were spaced 2·5 to 4 cm. apart and they were therefore chosen as sample-points, being both recognizable and slightly less far apart than the maximum permissible according to theory; thus there were 94 sample-points per run.

It can be seen from the autocorrelation diagram (Fig. 2) that the positions of successve course samples are correlated at about $r = -0{\cdot}4$. While the subject could in principle derive some help from this correlation, the calculation of input entropy has been based on the assumption that they were independent, i.e. that the entropy depended only on the probability distribution of sample values of the course. This distribution (Fig. 3) was roughly rectangular; with a grouping interval of $\frac{1}{2}$ cm. there were 16 groups in the maximum excursion of 7 cm. The input entropy was therefore approximately $\log_2 16$ or 4·00

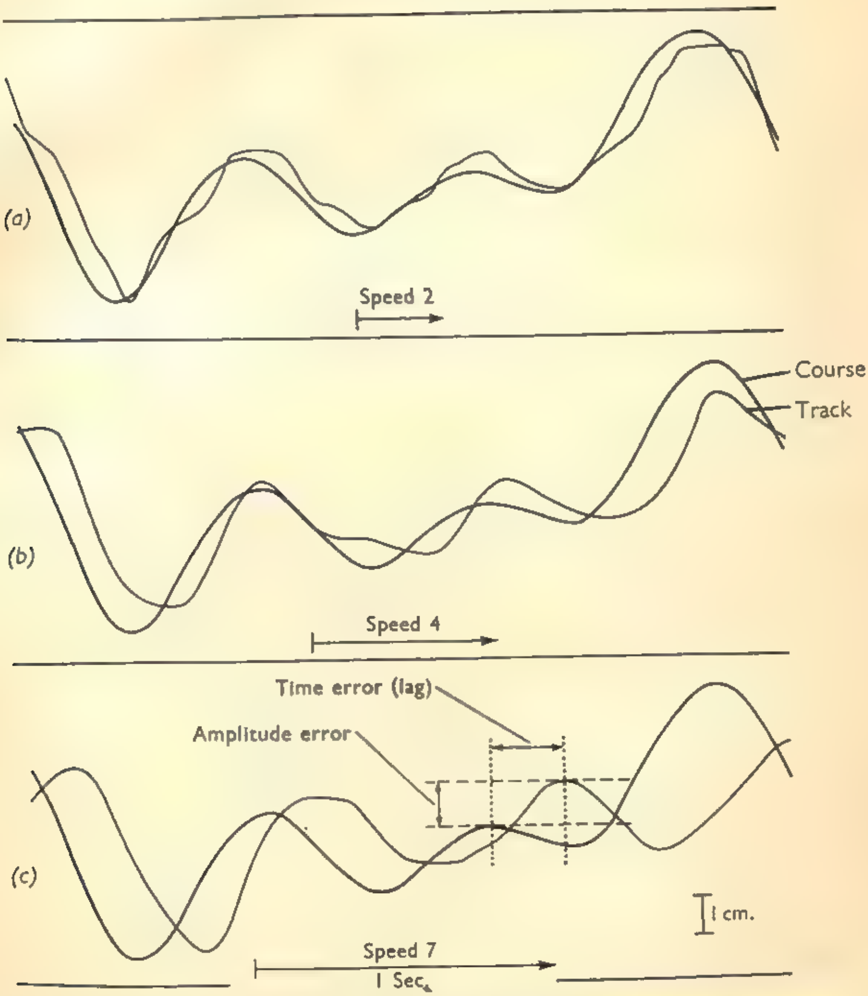
FIGURE 3

The amplitude distribution of course samples. (94 samples; grouping interval $\frac{1}{2}$ cm.).

bits per sample, though the value calculated from the actual distribution is 3·91 bits. The choice of $\frac{1}{2}$ cm. as grouping interval is equivalent to a choice of $\pm\frac{1}{4}$ cm. as the criterion of fidelity in performance, so that a maximum of 3·91 bits could be transmitted per sample if the subject were to keep his error to less than $\frac{1}{4}$ cm at each sample-point. For a less stringent criterion, the information per sample would decrease; for example, a unit of 1 cm. ($\pm\frac{1}{2}$ cm.) decreases the input entropy and the maximum rate to 2·91 bits per sample.

The track was found to lag behind the course by a fairly definite amount on any one run (see later discussion and Fig. 4). In principle, this lag represents a distortion which should be corrected by computing the correlation between course and track for various lags, and then adjusting the track by the one giving maximum correlation. (In so far as the lag fluctuates, it represents a calculable loss of information.) However, for each course peak the corresponding track peak was simply identified by visual inspection (see Fig. 4) and for each of the 94 pairs the course and track amplitudes i and o relative to the left-hand edge of the paper were recorded, the amplitude error e being obtained by subtraction. Very occasionally (once in a thousand peaks) it was impossible to find a proper sample point on the track and then the error was measured from the average value of the track.

FIGURE 4



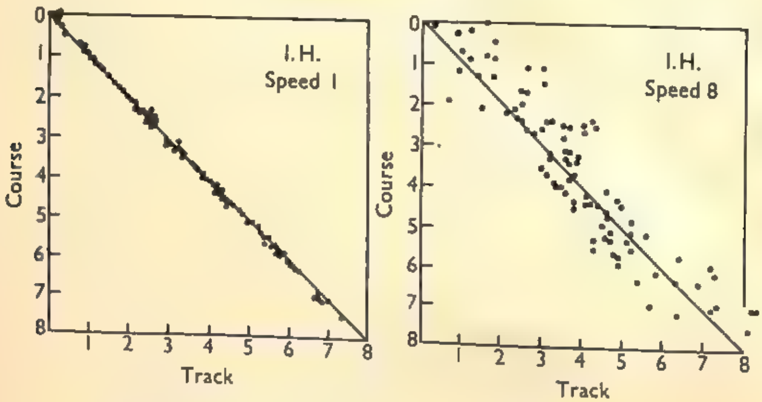
Examples of tracking performance and the identification of amplitude and time errors.

(*a*) Speed 2, no preview. (*b*) Speed 4, no preview. (*c*) Speed 7, no preview.

The 94 pairs obtained from a run were now entered in a scatter-diagram or input-output matrix (Fig. 5). A perfect performance would be represented on the diagram by a diagonal straight line, and complete failure either by a vertical straight line or by a random distribution of dots. The examples given illustrate the fact that performance was found to deteriorate in three ways with increasing speed. First, there was always more scatter

about the mean regression line (more random error or "noise"); second, the mean regression line sometimes had a steeper slope; third, and more rarely, the track was sometimes distributed bimodally, the regression line being then in two parts each parallel to the axis, suggesting that the subject could do no more than get his pointer onto the correct side of the mid-line. The first and third of these represent loss of information, but the second is a distortion, which was automatically allowed for in the later computations.

FIGURE 5



Scattergrams of course and track samples.
(94 samples are plotted; for perfect performance all points would lie on the diagonal).

There are several alternative formulae from which the input, output and joint distributions in the scattergram could be used to compute the entropies and the information-rate (see Shannon and Weaver, 1949, Chap. IV); that of equation (2), was used here. If a very large number of sample-points had been available on each run, the input-output matrix would have yielded functional probability-distributions, and the entropies could then have been evaluated with small error. But there were only 94 samples in each run and since the input sample-points were predetermined, not much improvement could be obtained by repeating runs. With a grouping interval of ½ cm. there were 15 possible sample-values of the course and track and so 15^2 cells in the input-output matrix. If all of them were occupied, the average cell-frequency would be about 0·5; but only 15 to 40 were actually occupied, giving an average frequency of 2 to 5. The estimated joint entropy $H'(i, o)$ was calculated from the cell-frequencies by the formula.

$$H'(i, o) = -\Sigma f(i, o) \log_2 . f(i, o) \qquad (4)$$

where $f(i, o)$ = relative frequency of entries in cell (i, o).

(To simplify the computation a table was drawn up to give the contribution of each cell-frequency to the total entropy, and the contributions were then added up directly.) The estimated entropies are perturbed both by sampling error and by systematic bias which are greater for smaller average cell-frequency. Draper (1954) and Hick (1956) have given expressions for them but it was not entirely clear how their formulae should apply in the present case; after trying different formulae the following correction for bias was chosen:—

$$H(i, o) = H'(i, o) - \frac{k - 1}{2N} \qquad (5)$$

where k = number of cells in the occupied region of the matrix
N = number of sample-points.

The correction term never exceeded 10 per cent. and was usually much less. Table I shows the result of applying a different grouping interval to the same data, with and without correction, from which it can be seen that the bias is partly but not completely offset by the correction. As the same interval (½ cm.) has been used throughout, comparisons between speeds and subjects should, however, be little affected by the residual bias in the present experiments.

TABLE I

THE EFFECT OF GROUPING INTERVAL ON THE ESTIMATED INFORMATION-RATE

(100 sample-points: Subject PW, Speed 4 with preview)

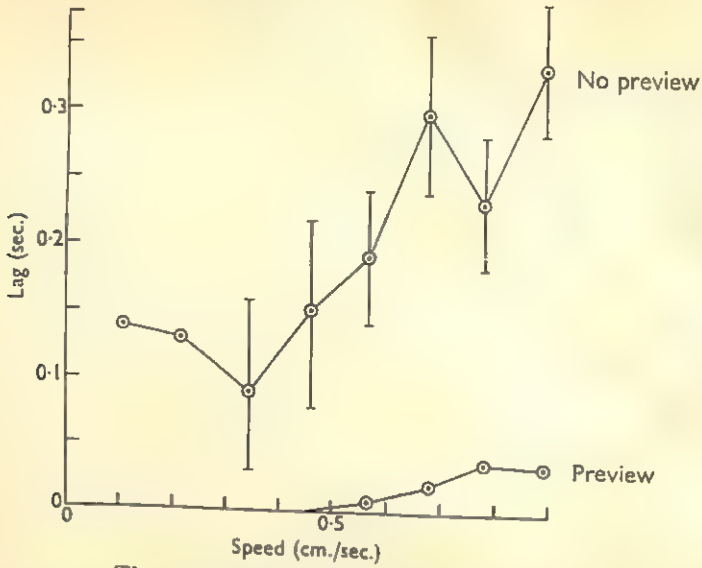
	Grouping Interval (*cm.*)		
	0·25	0·50	1·00
Input entropy	4·74	3·82	2·88
Cells occupied (k)	30	16	9
Corrected value	4·89	3·90	2·92
Output entropy	4·72	3·83	2·91
Cells occupied	29	16	9
Corrected value	4·86	3·91	2·99
Joint entropy..	4·84	4·88	3·45
Cells occupied	150	54	9
Corrected value	6·59	5·14	3·50
Rate per sample	3·62	2·77	2·35
Corrected value	3·16	2·67	2·41
Rate per sec.	5·25	4·02	3·41
Corrected value	4·58	3·87	3·49

RESULTS

At low speeds (Fig. 4(*a*)) subjects track accurately. They do not produce smooth tracks with the same bandwidth as the course, as they would were they acting as linear servo-mechanisms. Instead there are many small "ripples" which cross the course several times between sample points, and are presumably due (Hick, 1948) to the discontinuous functioning occasioned by the subject's finite reaction time. The bandwidth of the track being greater than that of the course, the track had to be "smoothed" before computing the rate, and this was done by inspection. At moderate speeds (Fig. 4(*b*)) the track is very similar to the course in general shape and follows it closely, usually with a small time lag, and some amplitude error; no discrete corrections can now be seen. At high speeds without preview (Fig. 4(*c*)) the track maintains the general shape and form of the course, but the lag becomes large and substantial errors of amplitude appear. One subject reported "I am doing no good at all; the thing is simply flashing by and I can do no more than make frantic attempts to catch up with it." In spite of this comment, the performance was considerably better than random, to the subject's later surprise. It was clear from the subjects' remarks that at speeds above 2 or 3 without preview, they aim at each successive peak as it comes up, and make no attempt to follow in between. Nevertheless, the track is a good copy of the course, which tends to confirm the rightness of the time-sampling procedure. With preview the deterioration is much less marked. Without preview the time-lag increased with speed up to a certain point (Fig. 6); with it there was a much smaller increase.

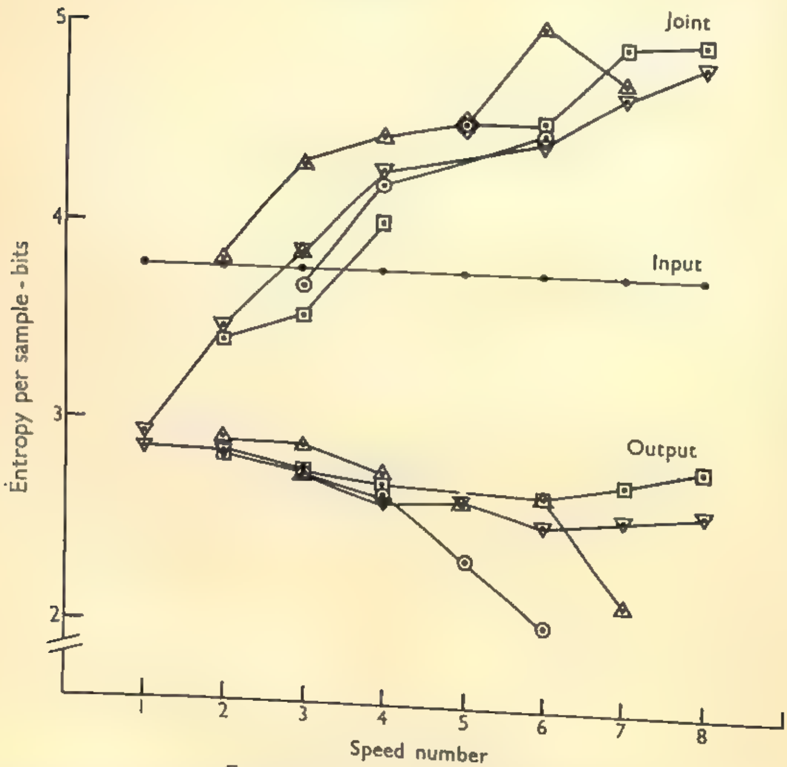
The output entropy per sample (Fig. 7) measures the scatter of the subjects' track or the amount of his movement. It showed a tendency to decrease at the higher speeds without preview, but not with preview. The joint entropy cannot be less than whichever is larger of the input and output entropies, it had this value at low speeds, and rose steadily as the speed increased which indicates a progressive loss of tracking accuracy by the subject, the rise being steeper without preview.

FIGURE 6



The average lag in tracking.
Each point is the average lag over one run for one subject.

FIGURE 7

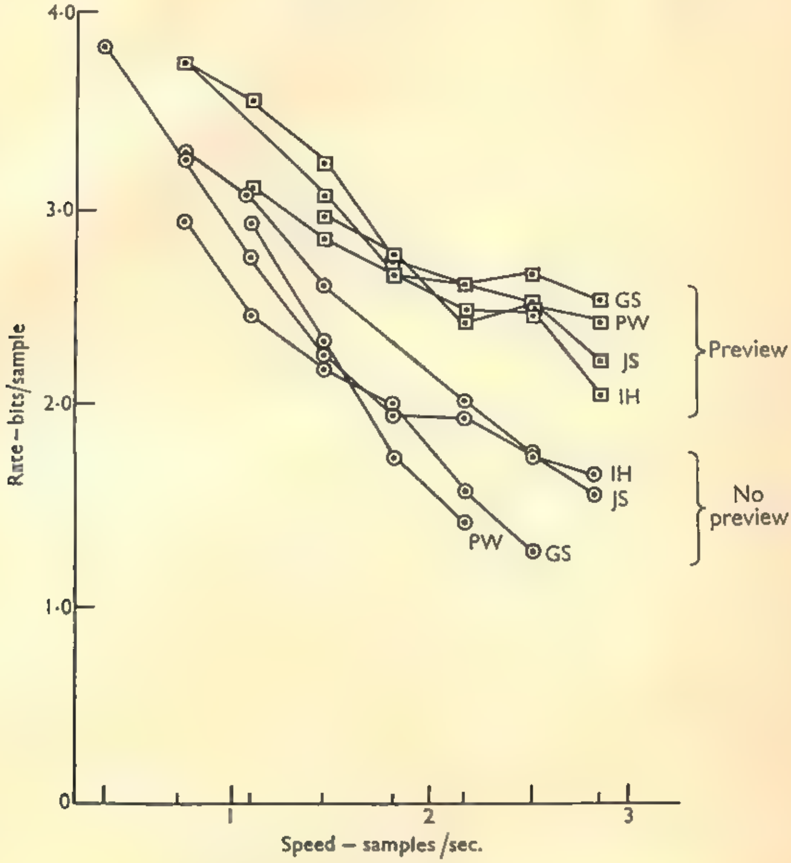


Course-speed and information.

Input, output and joint entropies are shown as a function of course speed. Each point represents one run for one subject. A correction for the sampling bias has been applied.

The information rate per sample (see Fig. 8) started equal to the input entropy, and fell steadily with speed. There was a clear difference between the preview and non-preview conditions at all speeds for all subjects, the preview condition giving the better performance.

FIGURE 8



Course-speed and information.

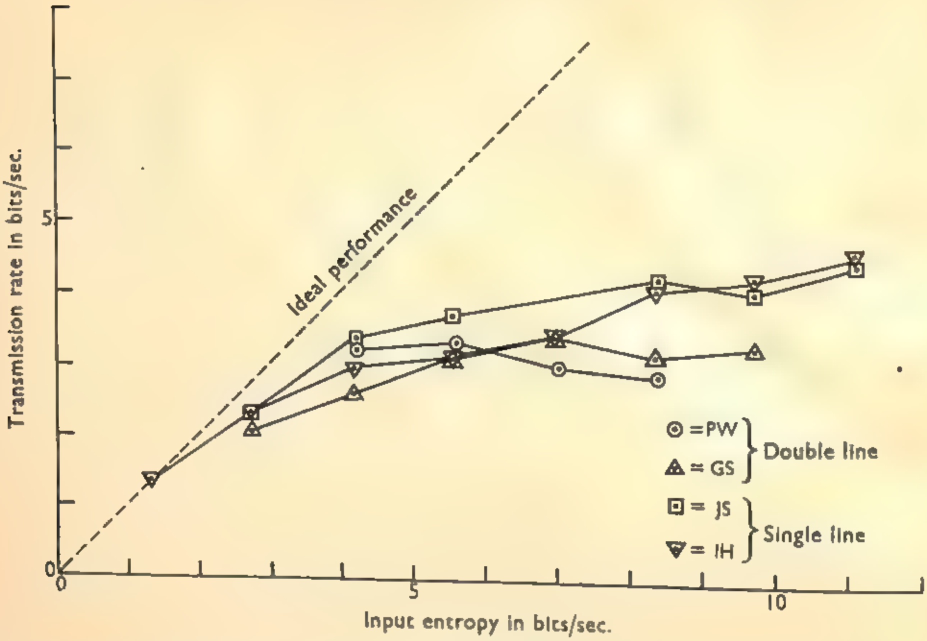
Information rate per sample point is shown as a function of course-speed. Each point represents one run by one subject.

The theoretical maximum information-rate per sec. is equal to the input-entropy, and hence the curve of ideal performance is a line with unit slope. All subjects achieved this at low speeds, beginning to depart from it at about 2 bits per sec. without preview and thereafter performance tended to level off at 3 to 5 bits per sec. (Fig. 9). The flat-topped curve suggests that the subjects had an informational "ceiling" or *channel-capacity* of about 4 bits per sec.

With preview the curve of performance began to depart from the ideal one at about 3 bits per sec. (see Fig. 10), and thereafter increased linearly. The average slope for all four subjects was about half (0·512 times) the ideal slope. At the highest input-entropy possible with the apparatus, there was no sign of the levelling off which would indicate the presence of a definite channel-capacity.

Practice might be expected to have some effect on the rate attainable for any given input speed, and so indeed it did. Table II shows the rates for three successive

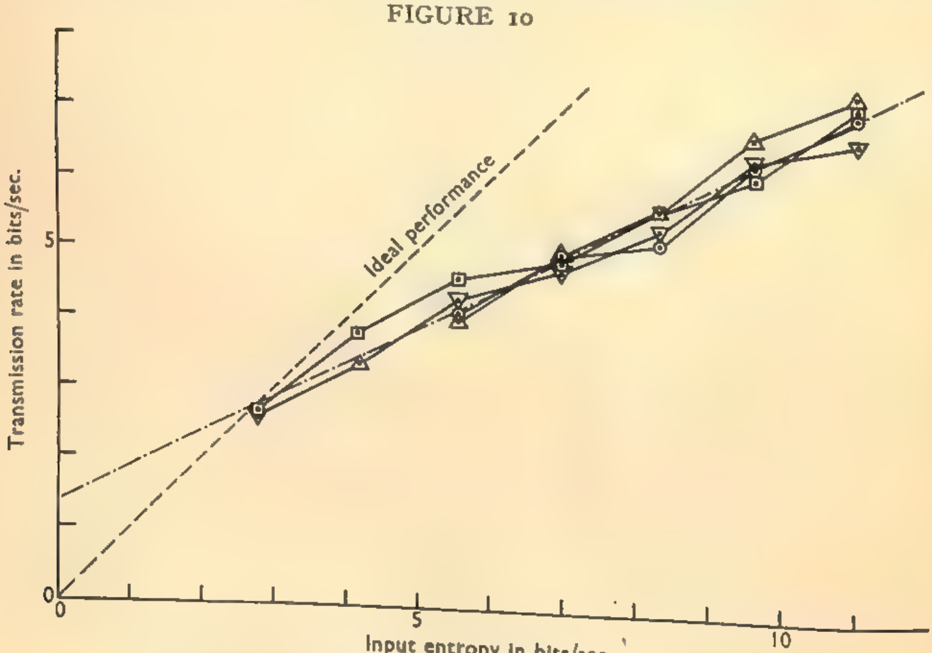
FIGURE 9



Course-speed and information—tracking without preview.

Information-rate per sec. is shown as a function of course-speed for pursuit tracking with minimal preview. Each point represents one run by one subject.

FIGURE 10



Course-speed and information—tracking with preview.

Information-rate per sec. is shown as a function of course-speed for pursuit tracking with ample preview. Each point represents one run by one subject. The dotted line has been fitted by the method of least squares; its equation is

$$R = 0{\cdot}51\ H(i) + 1{\cdot}36 \text{ bits}$$

runs for one subject on the same condition. There was a positive but relatively not large increase in information-rate. This is in line with the results of earlier studies of pursuit tracking. No attempt was made to measure fatigue effects, but within single runs at high speed there were indications from subjects' reports that fatigue was beginning to be felt.

TABLE II

THE EFFECT OF PRACTICE ON INFORMATION-RATE IN PURSUIT TRACKING WITHOUT PREVIEW

(Subject PW; Speed 4)

Run	*H(i)* *bits*	*H(o)* *bits*	*H(i, o)* *bits*	*Rate,* *bits per sec.*
1.	3·91	3·63	6·20	3·04
2.	3·91	3·74	6·20	3·16
3.	3·91	3·93	6·22	3·32

DISCUSSION

In recent studies Information theory has been used to analyse a number of sensori-motor tasks, and the human channel-capacity has been measured in two distinct ways. In discrete choice-tasks such as light-key reactions (Hick, 1952) and card-sorting when the subject does not know exactly what he will have to do on each trial—his uncertainty being measured by the *input entropy*—the average capacity seems to be about 5 bits per sec., though with very long practice, as in typewriting or piano-playing, (Quastler, 1956) or with high "compatibility" as in target-aiming (Crossman, 1956), he may reach 15 to 20 bits per sec. In these studies information rate may be measured by the total time taken to make a response of given information, or by the increment of time needed to deal with an increment of information, and the two approaches yield rather different results (Crossman and Szafran, 1956). In particular the time taken to complete the response, which is strictly irrelevant since it does not depend on information, has to be small and constant if the two methods of calculating are to give similar results.

The second type of channel-capacity, however, appears in the response itself, as a relation between speed and accuracy. Hand movements only have so far been subjected to study, and the pioneer work is that of Fitts (1954) confirmed by Crossman (1956) and Annett, Golby and Kay (1958). In Fitts' experiment the subject had to tap in turn two metal plates of width w set a distance a apart as fast as he could. The time per tap, t, was found to obey the relation—

$$t = -k \log\left(\frac{w}{a}\right) = -k\,(\log w - \log a) \qquad \cdot\cdot \qquad \cdot\cdot \qquad (6)$$

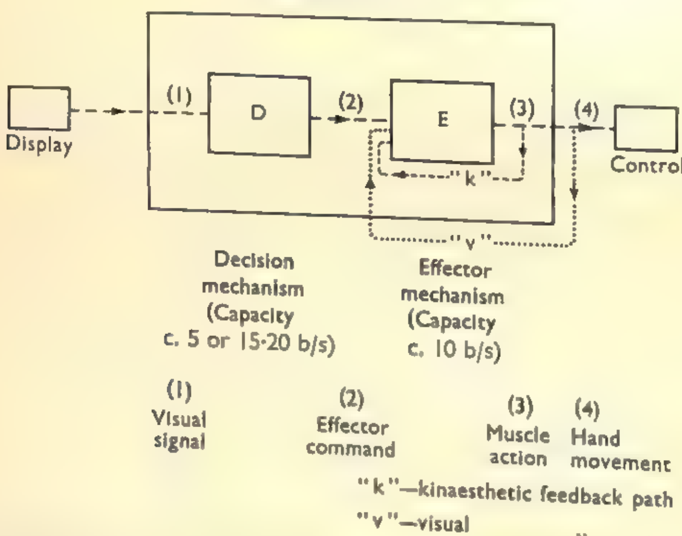
where k is a constant for one subject

a result which will be referred to as Fitts' Law. The right-hand side of the equation can be interpreted informationally (Crossman, 1956) as the difference between an initial entropy log a, and a final one log w, and hence it represents the reduction of uncertainty of the endpoint of movement achieved on a single tap. The information-capacity (the reciprocal of k) was found to lie between 8 and 15 bits per sec. in all the studies mentioned. In this task the subject knows exactly what to do on each trial, so the input entropy is zero. Therefore the time cannot, as in the previous case, be occupied in transmitting the information in the input, and Fitts suggests that it is

taken up in overcoming the "internal random noise" which prevents a subject repeating a movement precisely.

On the evidence of these studies it seems reasonable to postulate that the human sensori-motor apparatus comprises at least two functionally distinct parts (Fig. 11). The first, which may be called the Decision- or D-mechanism, is concerned with translating visual or other signals into orders which the second, the Effector- or E-mechanism, carries out. The D-mechanism has a capacity of 5 bits per sec. in most cases, and up to 15 or 20 bits per sec. in highly practiced or "compatible" tasks; the E-mechanism achieves about 10 bits per sec. for hand-movements. According to this model a subject's performance at a task such as pursuit tracking will depend on exactly how the two mechanisms are loaded and on the time-relations between their activities and the external situation.

FIGURE 11



A schematic representation of information-flow in the human perceptual-motor system.

The Decision (D-) mechanism translates perceived signals into instructions for the effector system.

The Effector (E-) mechanism controls the muscular activity needed to carry out the instructions. It may use visual feedback, "v," and/or kinaesthetic feedback "k."

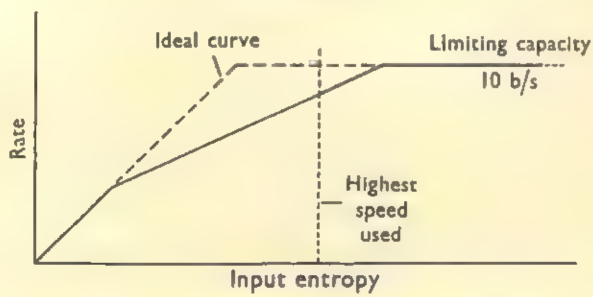
(a) *Tracking with preview*

The tracking task certainly had high "compatibility" and the D-mechanism would therefore be expected to work at 15 or 20 bits per sec.: though the control was by hand-wheel rather than direct movement, it seems reasonable to ascribe a capacity of about 10 bits per sec. to the E-mechanism. With preview the subject could look ahead as far as he liked and so offset the delays due to the working of the D- and E-mechanisms, by receptor anticipation (Poulton, 1957). This being so, the smaller of the two capacities should be limiting and one might thus expect an overall capacity of 10 bits per sec. This level was never attained in the experiment (see Fig. 10), though no definite ceiling appeared at a lower level.

The fact that the performance increased linearly with input load, yet at less than the ideal slope, suggests that another limitation, one allowing a constant fraction, rather than a constant absolute amount of information to be transmitted, must have been

present, and its locus may be sought in the division of visual attention between course and error. At high speeds, when looking ahead at the course, the subject was only able to see his own pen in peripheral vision, and actually to look at it would have entailed missing some of the course. Now the information-capacity measured by Fitts was for a fully visual task, represented in Figure 11 by the E-mechanism with the visual feedback path "v" in action. An equivalent task done without visual control has been studied by Vince (1948) who reports that hand movements of a few inches extent made without vision were distributed normally with about $\pm$ 7 per cent. error, and took 0·25 sec. or less. This much error is equivalent to $\log_2 100/14$ or about 2·8 bits per movement. In the present task, it would imply that as soon as the subject's visual channel, attending to the course some distance ahead of his pen, were loaded sufficiently near to its full capacity of 15 or 20 bits per sec. to be confined there, the information transmitted per *sample* should settle at about 2·8 bits, and the rate per sec. would be given by 2·8 $\times$ (number of samples per sec.), until the E-mechanism's upper limit of 10 bits per sec. were reached. The observed performance curve agrees reasonably well with this view, and it may be noted in its support that the curve of transmitted information (Fig. 8) does appear to flatten out at about 2·5 bits per sample. Figure 12 shows how the performance curve for the preview condition might be expected to behave at loads greater than those used in the present experiment. It is also noteworthy that subjects reported that they had adequate time for movement at even the highest speeds in the preview condition. If this explanation is correct, then at high speeds with preview subjects made purely positional responses to successive peaks and ignored the velocity and acceleration of the track in between.

FIGURE 12



The expected behaviour of subjects tracking at higher speeds than those used. The upper limit represents a channel-capacity of 10 bits per sec.

(*b*) *Tracking without preview*

Without preview the subject's attention must be centred on the narrow slit in which both course and controlled member appear. Bearing in mind that subjects were instructed to *reproduce* rather than strictly *track* the course, and that lag did not necessarily diminish their computed performance, the capacity expected according to the model would be just the same as in the preview condition, but in fact a capacity of only 4 bits per sec. was found. Again according to the model, the lag should be $\left(\frac{3{\cdot}91}{20} + \frac{3{\cdot}91}{10}\right)$ or about 0·6 sec., if the subject makes a positional response to each peak in turn; yet the observed lag (Fig. 6, Table III) was less than half this amount, which suggests that subjects may have been sacrificing information (i.e. accuracy of reproduction) to reduce their lag. Least lag is introduced by responding quickest, that

is at low accuracy, but for correct tracking each response must on average contain as much information as the track does in the same time. Therefore if the subject is trying to reduce his lag at low track speeds, one would expect to find him making frequent responses of low accuracy, and vice versa. In order to examine this point, one subject's records were examined closely; it was found that individual response movements could be distinguished sufficiently well to count them and measure their duration, with results given in Table III. At low speeds (1, 2 and 3) the expected behaviour clearly occurs; it also appears that even when allowance is made for the residual receptor anticipation permitted by the small preview, this subject must have been predicting course positions (perceptual anticipation) to some extent, since response movements were consistently initiated before their targets came into view. At higher speeds (4 and above) the number of movements fell to one per sample-point as had been expected, but the lag was still less than the movement-time at all speeds except possibly the highest, and the subject must therefore always have been anticipating course positions.

TABLE III

	Speed							
	1	2	3	4	5	6	7	8
(a) Characteristics of the tracking task								
Paper speed (cm. per sec.)	1·09	2·22	3·40	4·53	5·62	6·75	7·80	8.93
Time per sample (sec.)	2·78	1·37	0·89	0·67	0·54	0·45	0·39	0·34
Input entropy (bits per sec.)	1·41	2·86	4·38	5·84	7·24	8·69	10·00	11·50
Preview time (8 cm.) (sec.)	7·3	3·6	2·4	1·8	1·4	1·2	1·0	0·9
,, ,, (½ cm.) (sec.)	0·46	0·23	0·15	0·11	0·09	0·07	0·06	0·06
(b) Information Rates (average of four subjects)								
Per sample—with preview (bits)	3·86	3·66	3·26	2·92	2·67	2·48	2·50	2·40
,, without ,, ,,	3·86	3·18	2·83	2·35	1·91	1·66	1·61	1·77
Per sec.—with preview ,,	1·39	2·67	3·66	4·35	4·95	5·50	6·42	7·07
,, without ,, ,,	1·39	2·32	3·18	3·50	3·54	3·68	4·10	5·20
(c) Tracking without preview: an analysis of one subject's performance								
Average lag (sec.)	0·14	0·13	0·09	0·15	0·19	0·30	0·23	0·33
s.d. of lag (sec.)	—	—	0·07	0·07	0·05	0·06	0·05	0·05
Average duration of response movements (sec.)	0·45	0·32	0·30	0·33	0·45	0·44	0·38	0·33
No. of movements per sample-point	4–8	3–4	2–3	1–2	1–2	1	1	1
Information per movement (bits) (for perfect performance)	½–1	1–1·3	1·3–2	2·0 or 3·9	2·0 or 3·9	3·9	3·9	3·9

According to Poulton (1952, 1957) and others, prediction of the future position of a course is based mainly on its speed at the time of observation. If the subject were capable of making such a prediction with the same accuracy and speed as he can observe position itself, the changeover would entail no loss in information-rate, but this seems unlikely as the accuracy of prediction must inevitably diminish with its range, and prediction is also perceptually more complicated than direct observation of position. In terms of the model, then, one might suppose the prediction to be carried out by a third mechanism with a capacity of 4 bits per sec. approximately, whose

output would pass to the D-mechanism for translation into motor instruction: but a more parsimonious explanation is suggested by the similarity between the rate found here and that observed in choice-tasks of low compatibility (e.g. Hick 1952), which leads one to suppose that the prediction task is an "incompatible" one and that the capacity of 4 bits per sec. can simply be ascribed to the D-mechanism operating in its usual manner. This hypothesis could be tested by setting up a simple choice-task involving prediction and measuring the capacity directly.

If it is indeed the use of prediction which lowers the rate, a change might be expected to occur when the course speed no longer allows time for it, that is, when peaks recur at less than about 0·3 sec. The subject should then be forced into making pure positional responses lagging by just the spacing of one sample, and the rate should rise steadily with speed from 4 to 10 bits per sec. Unfortunately the highest track speed available fell just short of the point where this would begin to happen.

It remains to consider why the subject should avoid lagging by the time (about 0·6 sec.) needed for pure positional responses. Four reasons can be suggested: (1) that subjects could have done so but were given insufficient practice to overcome their initial reluctance to "get behindhand"; (2) that too great difference between the current motions of course and track would have produced intolerable visual confusion; (3) that course information might have been forgotten during such a comparatively long delay; (4) that too little feedback information about the response movements would then have been available. Further work will be needed to decide between these alternatives.

Conclusions

The informational analysis of pursuit-tracking performance appears to be feasible and to yield fresh understanding of the mechanisms involved. According to the relatively crude results of the present study the subject's performance seems capable of being fully explained in terms of a theoretical model derived from the results of previous work. In particular, the capacity in tracking with preview is just that of the effector system acting without visual feedback at about 2·5 bits per response or 10 bits per second, whichever is less. Without preview and when the lag is kept small the limit appears to be set at about 4 bits per sec. by the decision-mechanism predicting future course-positions.

It is hoped to carry out further work on these lines, with certain improvements. First, the present study was hindered by the need to use a predetermined course pattern, which had too much autocorrelation to be really satisfactory. It now appears technically feasible to provide a truly random course-generator of limited bandwidth. Second, it is possible that the speed of eye-movements may limit certain features of performance, such as the division of attention between course and track in the preview condition; eye-movements should therefore be recorded and compared with the current hand movements. Third, in the discussion free use has been made of information-capacities drawn from earlier work. Since there is considerable variation between individuals, a much more accurate analysis would be possible if each subject's capacity were measured on subsidiary tasks at the time of his tracking performance. Fourth, the present approach could also be extended to study the effect of statistical structure in the course, that is, how far the subject can make use of observed patterns of course behaviour to facilitate his performance.

Finally it may be said that the statistical information-theory approach to the study of tracking outlined in this paper is quite compatible with the analytic approach. It does, however, appear that the informational conditions of tracking should be

closely specified before a definite transfer-function can be identified by analytic means, and that its parameters may be expected to change with the statistical structure of the task.

The Author wishes especially to thank Mr. A. T. Welford and Dr. J. Szafran, who made the apparatus available and helped considerably in discussion. The apparatus was designed and built for studies of tracking performance by Mr. E. Earle, at the Nuffield Unit for Research into Problems of Ageing, Cambridge.

The research was carried out at the Department of Engineering Production, Birmingham University, under the general direction of Professor T. U. Matthew and later of Dr. N. A. Dudley and was sponsored by the Joint Committee on Individual Efficiency in Industry of the Medical Research Council and Department of Scientific and Industrial Research. It was financed from Counterpart Funds derived from United States economic aid.

References

Annett, J., Golby, C. W., and Kay, H. (1958). The measurement of elements in an assembly task—the information output of the human motor system. *Quart. J. exp. Psychol.*, **10**, 1–11.

Crossman, E. R. F. W. (1956). The measurement of perceptual load in manual operations. *Unpublished Ph.D. Thesis, Birmingham University.*

Crossman, E. R. F. W., and Szafran, J. (1956). Changes with age in the speed of information-intake and discrimination. *Experientia Supp. IV.* 128–35.

Draper, J. (1954). An introduction to information theory with reference to the human operator. Paper presented to conference at the Military College of Science, Shrivenham.

Fitts, P. M. (1954). The information-capacity of the human motor system in controlling the amplitude of hand-movements. *J. exp. Psychol.*, **47**, 381–91.

Fitts, P. M., Noble, M., and Warren, C. E. (1955). The frequency-response of skilled subjects in a pursuit-tracking task. *J. exp. Psychol.*, **49**, 249–56.

Hick, W. E. (1948). The discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, **1**, 36.

Hick, W. E. and Bates, J. A. V. (1951). The human operator of control mechanisms. *Ministry of Supply Monograph* 17, pp. 204.

Hick, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11–26.

Hick, W. E. (1956). Some miscellanea on information theory and the human operator. *Ministry of Supply Report WR/D/2/56.* Ed. Draper.

North, J. D. (1950). The human transfer-function in servo-systems. *Ministry of Supply Report WRD/6/50.*

Poulton, E. C. (1952). The basis of perceptual anticipation in tracking. *Brit. J. Psychol.*, **43**, 295–302.

Poulton, E. C. (1957). On prediction in skilled movements. *Psychol. Bull.*, **54**, 467–78.

Quastler, H. (1956). The force of habit. *Univ. of Illinois Control Systems Laboratory Report R-70.*

Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication*, Urbana, pp. 117.

Vince, M. A. (1948). Corrective movements in a pursuit task. *Medical Research Council A.P.U. Report No.* 75.

CONCERNING THE INTERRELATION BETWEEN ABSOLUTE SENSITIVITY AND STRENGTH OF THE NERVOUS SYSTEM

BY

V. D. NEBYLITSYN, V. I. ROZHDESTVENSKAYA and B. M. TEPLOV

From the Institute of Psychology, Moscow

SUMMARY*

This paper is concerned with the relation between absolute sensitivity and the "strength" of the C.N.S., in the sense in which the term was used by Pavlov. The authors advance the hypothesis that "weakness" of the nervous system is correlated with high reactivity, and hence with high sensitivity to peripheral stimulation; conversely, "strength" of the nervous system is correlated with low reactivity, and hence with low sensitivity to peripheral stimulation.

Two sets of experiments were carried out in order to test this hypothesis, one concerned with visual and the other with auditory sensitivity. In the visual experiments, absolute sensitivity to light in the periphery of the dark adapted eye was measured on 26 subjects by means of an adaptometer. Measurements were taken at two minute intervals over a period of about half an hour, the mean threshold value for each subject being calculated from the results obtained at a number of experimental sessions distributed over 3 to 5 days. In the auditory experiments, absolute threshold for sound intensity was determined in 22 subjects. A pure tone of 1,000 c.p.s. was fed to the subject binaurally through headphones and the sound pressure level systematically varied by the experimenter. Threshold measurements were made at 1 min. intervals over a 15 to 20 min. period. Sensitivity is calculated as a value inversely proportional to that of the threshold.

In order to determine the "strength" of nerve cells, the authors made use of a method involving a conditioned decrease of sensitivity to light or sound. The rationale of the method is as follows: The strength of a conditioned reflex is normally found to increase with increase in intensity (or duration) of the conditioned stimulus until a "physiological limit" is reached. Above this limit, further increase in stimulus intensity no longer produces an increase in response, which may indeed diminish as a consequence of "transmarginal inhibition." Hence a determination of the stimulus intensity above which there is no increase in response strength provides a method of defining the physiological limit, or "strength," of the central nervous processes involved.

The actual method used here consisted in exposure for 10 to 15 sec. of an illuminated panel as the unconditioned stimulus, the resulting reduction in peripheral visual sensitivity being the unconditioned response. A pure tone of 1,000 c.p.s. and a red light of low intensity were used as conditioned stimuli. The conditioned stimulus was applied 5 sec. before the onset of the unconditioned stimulus, the duration of their combined application being 10 sec. After the conditioned reflex was well established the tests of "strength" were begun. These consisted in the application of the conditioned stimulus (always reinforced) ten times in succession at intervals of 2 min. The conditioned stimulus was also presented twice without reinforcement, once at the beginning and again at the end of the series. If the strength of the conditioned reflex showed no decrement on this second trial as compared with the first, it was concluded that the physiological limit or "strength" of the corresponding central nervous structures was high; if, however, there was decrement, this was taken to mean a low physiological limit or "weakness." The criterion of decrement was a decrease in strength of the conditioned reflex of 10 per cent. or more.

These experiments were also repeated after the administration of 0·2 gr. of caffeine, which is thought to raise the excitability, and hence diminish the "strength," of central nervous processes.

* This summary has been prepared by the Editor. He wishes to thank Professor Teplov for his authority to make certain changes in the text, and Professor A. R. Luria and Dr H. B. Barlow for valuable help in the preparation of this paper.

The results appear to demonstrate a negative correlation between sensitivity, visual or auditory, and the "strength" of nervous processes. In general, subjects with "weak" nervous systems gave low absolute threshold values, indicating high sensitivity, whereas those with "strong" nervous systems gave high threshold values, indicating low sensitivity. There were, however, some exceptions to this rule.

In discussion of these results, the authors lay stress on functional changes which are believed to occur in neurones as a consequence of sustained excitation. In general, they argue that both sensitivity and "strength" of nervous processes can be related to a more basic dimension of neural "reactivity."

I. Introduction

Pavlov and his collaborators showed that the strength of the nervous system is one of the most important properties which determines individual differences of the higher nervous activity in animals and man. The parameter of the strength of the nervous system was taken by Pavlov as a basis for his classification of the types of nervous system.

According to Pavlov, the strength of the nervous system is characterized mainly by the limit of the working capacity of the nerve cells, i.e. by their capacity to endure a protracted and concentrated excitation or action of a very strong stimulus, without passing into the state of inhibition. As to the weakness of the nervous system, this is expressed in a low working capacity or a low limit of endurance of the nerve cells. Externally this is manifested in the emergence of transmarginal inhibition when the intensity or duration of the action of the stimulus is increased beyond a definite limit.

Thus, from the point of view of the endurance of the cells the weakness of the nervous system is a negative property. Taking only this aspect into consideration one may, naturally, regard the weak type of nervous system as a "bad" one, and the strong type as a "good" one. Such a viewpoint appears almost self-evident. Pavlov himself adhered to it, characterizing the weak type of nervous system as a "more or less invalid" type.

This viewpoint, however, seems to us one-sided and consequently erroneous.

As shown by the analysis of life experience, persons with a weak nervous system, i.e. a nervous system characterized by low endurance, easy fatiguability, etc., exhibit a number of peculiar features showing that their nervous system possesses not only negative but also manifestly positive properties.

This is confirmed by some of the theoretical propositions of Pavlov himself. In his works written in the period from 1922 to 1927 he repeatedly expressed the idea that the rapid exhaustibility of weak nerve cells is closely connected with their "extreme irritability," or "easy excitability" (Pavlov, 1926).

These facts and theoretical propositions, which we expounded in greater detail elsewhere (Teplov, 1956), have led us to the hypothesis that *the weakness of the nervous system is internally connected with its high reactivity, or high sensitivity.*

The present work is devoted to the experimental corroboration of this hypothesis.

II. Methods of Investigation

The principal method used by us was the method of comparing experimental data obtained from one and the same person and relating, on the one hand, to absolute sensitivity, and, on the other hand, to the strength of the nervous system.

1. *Determination of sensitivity*

(*a*) *Visual sensitivity.* The measurements of the sensitivity of peripheral vision are carried out with the help of an adaptometer. The subject is placed in a dark room. The white circle of the adaptometer (having a diameter of 1°30′) is fixed on the wall in front of the subject. At the side of the circle (angular distance—10°) is a red fixation dot which serves for fixing the eyes of the subject during the measurement of sensitivity. A ray of light from the adaptometer (placed behind the subject) falls upon the circle. The

ray is intersected by an absorbing photo-filter. Moving the latter from one place to another, the experimenter is able gradually to change the degree of illumination of the circle. The experiment is preceded by a dark-adaptation during 45 min. At the signal "attention!" the subject fixes his eyes on the red dot. He has to ascertain, as precisely as possible, the moments of emergence and disappearance of a just perceptible bright spot in the place of the circle and immediately to inform the experimenter. The latter fixes the corresponding divisions on the photo-filter which indicate the percentage of the absorption of light. The mean of these two values characterizes the threshold of visual sensation. The sensitivity is calculated as a value which is inversely proportional to that of the threshold. Measurements are taken at intervals of 2 min. during a period from 20 to 40 min., until the indications of the subject acquire a uniform and stable character. The average value is deduced on the basis of the results obtained during 3 to 5 experimental days.

(*b*) *Auditory sensitivity*. The measurement of auditory sensitivity is carried out with the help of a sound generator in a room relatively insulated from outside noise. The sound (a pure tone of 1,000 c.p.s.) is conveyed to the subject binaurally through ear-phones. Smoothly turning the handle of the rheostat, the experimenter gradually modifies the output voltage and, consequently, the level of sound pressure in the acoustic duct of the subject, first increasing it and then decreasing. The subject must ascertain the moments of emergence and disappearance of the sound and immediately inform the experimenter. The latter fixes corresponding divisions on the scale which is graduated in units of sound pressure. The mean of these two values characterizes the threshold of auditory sensation. The sensitivity is calculated as a value which is inversely proportional to that of the threshold. Measurements are taken at intervals of 1 min. during a period from 15 to 20 min., until the indications of the subject reach stability.

2. *Determination of the strength of nerve cells*

A special method of determining the strength of nerve cells has been elaborated. It was necessary to establish the individual limits to which concentrated excitation in the nerve cells can be increased. According to the so-called "law of force," the strength of a conditioned reflex increases with the growing intensity of the conditioned stimulus, all other things being equal. If, however, the intensity of the stimulus oversteps the limits of the working capacity of the nerve cells, the strength of the conditioned reflex no longer increases; on the contrary, it diminishes, since "transmarginal" or "protective" inhibition comes into force. A similar effect is obtained if we increase the duration of the stimulus and not its intensity; in this case the stimulus may act not continuously, but with interruptions. It is only important that this stimulus should evoke well-concentrated excitation in the cerebral cortex; excitation in the focus of the conditioned stimulus as a result of a firmly elaborated conditioned reflex fully complies with this requirement.

In our experiments, we used a conditioned decrease of sensitivity to light or sound. The unconditioned stimulus consisted in a momentary (10 to 15 sec.) illumination of a panel placed before the subject and occupying his entire field of vision. The intensity of illumination at the subject's eye-level was about 40 lux. As a result of this stimulus, sensitivity of peripheral vision was as a rule reduced by some 50 to 80 per cent.; in 2 to 5 min., its sensitivity was restored to its previous level.* A pure tone of 1,000 c.p.s. conveyed to the subject from a sound generator through ear-phones and a weak optic stimulus of red colour (which does not disadapt the rods of the retina) were used as conditioned stimuli. In some cases the very first presentations of these stimuli produced a small increase or decline of sensitivity (by 10 to 20 per cent). These changes in sensitivity were, apparently, manifestations of a general orienting reaction, since they were readily extinguished in the course of one or two experiments. After the extinction of the orienting reactions we combined the conditioned stimulus which acted during 15 sec. with a 10-sec. bright illumination, the conditioned stimulus being put into action 5 sec. earlier than the unconditioned stimulus. After a series of such combinations a conditioned connection was established between the action of the conditioned stimulus and the reaction of a decline of visual sensitivity. This was shown by the fact that the conditioned stimulus, when applied alone, began to cause a similar decline of sensitivity. This

* The rôle of the pupil changes in this conditioned reflex has been specially investigated by V. I. Rozhdestvenskaya (1955), who was able to show that a conditioned decrease of sensitivity to light is also observed under conditions in which pupil size is maintained constant by means of homatropine. Under such conditions, however, the strength of the reflex declines considerably.

decline in general amounted to 15 to 30 per cent. of the background level of sensitivity, i.e. it does not reach the strength of the unconditioned reflex, which usually lies within the range of 50 to 80 per cent. The conditioned decline can still be observed 30 to 60 sec. after the termination of the conditioned stimulus.

After the elaboration of a well-established conditioned reflex (which could be seen from the occurrence of this reflex in not less than three successive experiments) we performed our first control experiment of testing the "strength" of the nervous system. In this experiment the conditioned stimulus, now calling forth a concentrated excitation (every time accompanied by a reinforcement in order to prevent the extinction of the conditioned connection), was presented ten times in succession at intervals of 2 min. Thus, concentrated excitation was produced in the same cells repeatedly. We assessed the state of the cells by the strength of the conditioned reflex. For this purpose, the conditioned stimulus was twice presented in isolation, i.e. without reinforcement: the first time at the very beginning of the experiment, and the second time immediately after its tenth presentation in combination with a reinforcement. If the strength of the conditioned reflex in the second test remained on the same level as in the first test, this was taken to indicate the high working capacity, or "strength" of the nerve cells. On the other hand, a decline of the strength of the reflex was taken to indicate a low working capacity, or "weakness" of the nerve cells. The cause of this decline lies in transmarginal inhibition arising in the cells whose low working capacity prevents the further growth of excitation within them.

The second control experiment on testing the "strength" of the nervous system was performed after the administration of 0·2 gr. of pure caffeine. Caffeine raises the excitability of the nerve cells, as a result of which the limit of their working capacity is reached at a smaller intensity or duration of the action of the stimulus. As was to be expected, after the administration of caffeine the decline in the strength of the conditioned reflex, as a result of its frequent repetition in combination with a reinforcement, was of a more pronounced character. In a number of cases the strength of the reflex remained unchanged in experiments not accompanied by the administration of caffeine, while in experiments where caffeine was applied we noted a decline. In these cases we have, apparently, an intermediate level of strength of the nervous system.

Using the terminology accepted in Pavlov's laboratories, we can define the aforementioned methods as *methods of extinction with the reinforcement of a "photochemical" conditioned reflex.*

In an earlier study (Nebylitsyn, 1957) we established that tests of the strength of the nervous system by the method described indicate primarily the strength of the nerve cells in the region of the brain to which the conditioned stimulus is addressed. Indeed in some subjects, extinction trials accompanied by reinforcement yielded different results according to the modality of the conditioned stimulus. We therefore endeavoured to determine the strength of the cells in the visual cortex by experiments using a conditioned visual stimulus, and the strength of the cells in the auditory cortex by experiments using a conditioned acoustic stimulus.

3. *Subjects*

The subjects were mainly students, of both sexes, with an age-range from 18 to 30. Four older subjects, with ages ranging from 43 to 55, were also employed. Before the extinction trials all subjects were thoroughly trained in making observations, the accuracy of which were specially checked by the experimenter.

III. Results

A. *Determination of visual sensitivity and of the strength of the cells of the visual cortex*

The results of this series of experiments are shown in Table I. It will be seen that the range of individual differences in sensitivity is very considerable. Thus the sensitivity of subject No. 1 in Table I is almost tenfold that of subject No. 26. These variations are in keeping with the data obtained by other authors. It will also be noted that, within the group as a whole, the numerical values of sensitivity form a continuous series.

When determining the strength of nerve cells, the criterion employed is a decrease in the strength of the conditioned reflex by 10 per cent. or more. This decrease is measured, not from the initial strength of the conditioned reflex, but from the

"background level" of sensitivity measured after the extinction trial accompanied by reinforcement. It is plain, however, that we cannot by such a method obtain gradations in the strength of nerve cells comparable to those of sensitivity. Our present methods allow us only to divide the subjects into two groups, those with weak and those with strong nerve cells respectively.

TABLE I

Absolute Visual Sensitivity in comparison with the Results of Extinction Trials accompanied by Reinforcement

Subjects	Absolute sensitivity (in conventional units)	*Strength of conditioned reaction in percentages of decline (or increase) of sensitivity*				Strength of nerve cells
		Without caffeine		*Caffeine 0·2*		
		Before extinction	*After extinction*	*Before extinction*	*After extinction*	
1	24·5	−26	−16	−26	+6	weak
2	23·3	−27	0	−30	+12	weak
3	22·1	−40	−18	−22	+5	weak
4	19·6	−32	−13	−24	0	weak
5	18·4	−45	−53	−29	−39	strong
6	18·4	−26	−5	−15	−7	weak
7	17·2	−28	−15	−40	0	weak
8	15·9	−21	+5	−27	+35	weak
9	15·9	−27	+14	−18	+18	weak
10	15·0	−38	−6	−15	+9	weak
11	15·0	−30	−11	−30	0	weak
12	13·4	−25	−5	−40	−5	weak
13	12·7	−19	−46	−38	−22	intermediate
14	10·2	−20	−13	−26	−18	strong
15	8·9	−26	−23	−24	−40	strong
16	8·9	−23	−49	−27	−46	strong
17	8·9	−15	−49	−17	−38	strong
18	8·5	−38	−38	−39	−38	strong
19	8·1	−20	−32	−26	−28	strong
20	6·7	−20	−35	−30	−37	strong
21	6·7	−21	−42	−29	−41	strong
22	6·3	−30	−13	−26	−6	weak
23	6·0	−27	−21	−20	−17	strong
24	6·0	−23	−36	−27	−60	strong
25	5·6	−20	−25	−26	−20	strong
26	2·5	−27	−38	−22	−15	strong

The first group comprises 12 subjects (cf. Table I). In two of these (Nos. 8 and 9), the conditioned stimulus, after the extinction trial accompanied by reinforcement, not only failed to evoke the usual decline but actually led to an increase in sensitivity. This increase is still more pronounced in experiments with administration of caffeine. The experiments of Demirchoglyan and Allakverdyan (1957) have already shown that inhibition, which develops when a conditioned "photochemical" reflex is extinguished without reinforcement, is accompanied by an increase in sensitivity. Similar reactions of an opposite sign have been recorded by Ilyina (1958) in experiments involving the application of differential stimuli. These findings give us grounds to believe that an increase in sensitivity after an extinction trial with reinforcement (as in the present experiments) is due to transmarginal inhibition of a more profound

character than in the case of a simple decline in the strength of the reaction. It must be assumed that individuals showing this phenomenon possess the weakest nerve cells.

The second group comprises 13 subjects (cf. Table I). In these individuals the strength of the conditioned reflex does not decline after extinction trials accompanied by reinforcement—indeed in some cases (Nos. 15, 16, 17, 21 and 24) it actually increases. This testifies to the fact that a further intensification of excitation in the relevant cortical cells is still possible; that is to say, the limit of their working capacity has not yet been reached. These subjects may be regarded as having the strongest nerve cells in the visual cortex.

The results in one subject (Table I, No. 13) were somewhat anomalous. In this case, although the strength of the conditioned reflex after an extinction trial under ordinary conditions did not increase, it declined considerably following the administration of caffeine. This suggests that the strength of the relevant nerve cells in this subject is at an intermediate level.

Let us now consider the distribution of these data with regard to the measurements of absolute sensitivity. It will be seen from Table I, in which subjects are enumerated in order of diminishing sensitivity, that 11 of the 12 subjects with weak nerve cells occupy the top 12 places in the table while 11 out of the 13 subjects with strong nerve cells occupy the bottom 12 places in the table. It would therefore appear that the subjects with the highest visual sensitivity are almost without exception those possessing weak cells in the visual cortex: and conversely, that subjects showing a relatively lower absolute visual sensitivity level are in the overwhelming majority of cases those possessing strong cells in the visual cortex.

These results give us ground to state that, in general, there is concurrence between high visual sensitivity and a weak visual cortex and between low visual sensitivity and a strong visual cortex. There are, however, two exceptions to this rule (Table I, Nos. 5 and 22). Further, we did not observe any gradual interdependence between sensitivity and strength within each of the basic groups. These two subjects (Nos. 8 and 9), who according to our criteria possessed the weakest nerve cells, did not reveal the highest sensitivity, and five subjects (Nos. 15, 16, 17, 21 and 24) regarded by us as possessing the strongest nerve cells by no means exhibited the lowest sensitivity. Further, the two subjects (Nos. 25 and 26) who stand at the bottom of Table I did not exhibit any signs of maximal strength. Indeed, in the caffeine experiments these subjects not only fail to manifest any increase in the strength of the reflex but even show a slight decline. In view of these considerations, it follows that the relation between sensitivity and strength is not of a functional nature but must be regarded simply as a correlation apparent when the group is considered as a whole.

A statistical analysis of the data by means of the Fisher z-test (Snedecor F test) revealed that the difference between the arithmetical means of the visual sensitivity thresholds in the two groups was highly significant ($F = 27.59$; $p < 0.001$, the degrees of freedom being 1 and 23).

The results of this series of experiments fully confirm our hypothesis concerning the co-ordination between sensitivity and strength of the nerve cells. High visual sensitivity is combined with weakness, and low sensitivity with strength, of nerve cells in the corresponding areas of the brain.

B. Determination of auditory sensitivity and of the strength of the cells of the auditory cortex

The results of this series of experiments are presented in Table II. As with the visual experiments, we note a wide range of individual differences in auditory sensitivity—the value given by subject No. 1 in Table II being over five times greater than that given by No. 22.

TABLE II

Absolute Auditory Sensitivity in comparison with the Results of Extinction Trials accompanied by Reinforcement

Subjects	Absolute sensitivity (in conventional units)	*Strength of conditioned reaction in percentages of decline (or increase) of sensitivity* — Without caffeine		Caffeine 0·2		Strength of nerve cells
		Before extinction	*After extinction*	*Before extinction*	*After extinction*	
1	27·0	−20	+14	−25	0	weak
2	26·0	−39	+ 7	−41	+ 3	weak
3	21·1	−22	+ 6	−25	0	weak
4	20·1	−28	0	−35	− 5	weak
5	18·0	−28	−10	−29	−10	weak
6	18·0	−25	−22	−18	0	intermediate
7	14·2	−28	−39	−28	−13	intermediate
8	13·2	−28	−38	−17	−34	strong
9	12·9	−32	0	−29	0	weak
10	12·1	−36	−24	−23	− 5	intermediate
11	11·7	−23	−41	−26	−34	strong
12	11·7	−40	−12	−23	− 7	weak
13	10·7	−29	+10	−34	0	weak
14	9·7	−17	−16	−30	− 6	intermediate
15	9·6	−22	−33	−22	−15	strong
16	9·1	−21	−26	−19	−46	strong
17	8·2	−28	−26	−23	−30	strong
18	7·8	−23	−21	−22	−31	strong
19	6·3	−28	−50	−20	−32	strong
20	6·2	−24	−19	−26	−20	strong
21	6·2	−23	−34	−16	−10	strong
22	5·0	−26	−32	−23	−17	strong

These data permit us to classify the subjects into three groups according to relative cerebral strength, viz. those with weak nerve cells (Group 1), those with strong nerve cells (Group 2), and an intermediate group (Group 3) (cf. Table II).

Group 1 comprises eight subjects, in all of which a decline in strength of the conditioned reflex following an extinction trial accompanied by reinforcement was observed. Group 2 comprises 10 subjects and in these, the strength of the conditioned reflex did not change following an extinction trial accompanied by reinforcement. Group 3 comprises four subjects in which the results were intermediate between those of Groups 1 and 2. In these four subjects, the strength of the conditioned reflex declined only after the administration of caffeine.

As regards the distribution of the subjects in these three groups with regard to auditory sensitivity level, the picture is somewhat more complicated than in the case of the visual experiments. It is true that the top five places in Table II are occupied by members of Group 1, but the other three subjects in this group occupy the 9th, 12th and 13th places respectively. Thus five subjects with a weak auditory cortex possess high sensitivity and the remaining three medium sensitivity. In this same way, the bottom eight places in Table II are occupied by members of Group 2, whereas

the remaining two subjects in this group occupy the 8th and 11th places respectively. Thus, while the majority of subjects with strong nerve cells possess relatively low auditory sensitivity, two show medium sensitivity. All four subjects in Group 3 will be seen to possess medium sensitivity.

Taking the picture as a whole, it will be seen that whereas the subjects at the two ends of Table II (viz. Nos. 1 to 5 and 15 to 22) show uniform level of strength, there is a zone (viz. Nos. 6 to 14) which includes all three levels of strength as revealed by our method of assessment. Thus we can state that the conformity between absolute sensitivity and strength of the nerve cells does not clearly hold in the case of the intermediate threshold values. But this conformity increases steadily as we approach either the higher or the lower threshold values for absolute sensitivity.

Statistical analysis revealed that the difference between the mean auditory sensitivity values in Groups 1 and 2 was highly significant ($F = 21{\cdot}42$. $p<0{\cdot}001$; the degrees of freedom being 1 and 16).

The results of these experiments provide confirmation of our hypothesis regarding the correlation between sensitivity and strength of the nervous system—weak nerve cells corresponding to higher sensitivity and strong nerve cells to lower sensitivity.

IV. Discussion

The data obtained in these experimental investigations of two sensory modalities have confirmed the hypothesis of the existence of a negative correlation between sensitivity and strength of the nervous system. We have now to elucidate the nature of the relationship between these two seemingly different parameters of nervous activity.

The concept of strength of the nervous system can be defined in terms of the functional stability of nervous tissue when subjected to conditions of stimulation which, from the standpoint of intensity or duration—or both—approach or even exceed the normal physiological limit. When, therefore, we speak of the strength of the nervous system as a typological property we have in mind the limit of working capacity as determined by the threshold of functional endurance: the higher this threshold, the greater the strength of the system.

We shall approach the concept of sensitivity from the standpoint of the internal physiological conditions which determine the excitability of the central nervous mechanisms governing sensory response. The term "threshold of excitability" (and in consequence "threshold of sensitivity") usually implies that the energy of stimulation, upon reaching a definite threshold value, evokes a state of excitation in the receptor mechanism. Hence thresholds of sensitivity, as psychophysically conceived, are primarily dependant upon excitability thresholds in the receptor and central mechanisms. The question of the nature of sensitivity thus resolves itself into a question of the nature of the basic excitatory process.

In 1940, Nasonov developed a theory of excitation based upon the concept of "paranecrosis" which has not become generally known outside its country of origin. Nasonov and his collaborators claim that cells react to damage with a complex of structural changes ("paranecrosis"), believed to be based upon the so-called denaturation of cell proteins. It is possible that changes of this nature may also underlie neural excitatory processes in living tissue (Nasonov and Aleksandrov, 1940), as is suggested more particularly by the work of Mirsky (1936) on retinal photochemistry. One may therefore suggest that the process of excitation in any biological system involves complex reversible changes at the level of microstructure.

If excitation in neurones actually involves alteration in cellular structure, and if thresholds of sensitivity relate to thresholds of neural excitability, we can regard sensitivity thresholds as points of transition between states of physiological rest and states of reversible

structural modification, i.e. as thresholds of reversible functional change. Hence sensitivity, interpreted as the capacity to perceive an external stimulus, is to be viewed as the outcome of a process in which neural activity is intensified through a reversible destruction of its own microstructure.

In the light of these considerations, we wish to suggest that both sensitivity and neuronal strength may be subsumed under the more general concept of "reactivity." This property may be said to characterize the integral function of any living tissue, being defined at one extreme in terms of its threshold of excitability and at the other in terms of its limit of effective working capacity (Pavlov's "limit of excitation"). "Reactivity" thus denotes the general function of an excitable substratum; it determines the course of its reaction for any adequate stimuli, beginning with the minimal response at threshold level and ending with the maximal strain on working capacity which is followed by a phase of protective inhibition. The concept thus embraces both the sensitivity of the system and its endurance in the face of adverse conditions of stimulation.

V. Conclusion

This paper is concerned with the inter-relationship between the strength and the sensitivity of the nervous system. Comparative data are adduced relating on the one hand, to levels of auditory and visual sensitivity and, on the other, to the strength of nerve cells (in Pavlov's sense of the term) in corresponding regions of the central nervous system. Strength, as so defined, has been ascertained by the use of a method involving a conditioned decrease in sensitivity to light or sound. A negative correlation has been found between the characteristics of sensitivity and strength of the nervous substratum; thus high sensitivity is combined with weakness, and low sensitivity with strength, of the underlying nervous processes. In conclusion, the authors attempt to relate the parameters of strength and sensitivity in terms of a single typological property.

References

Demirchoglyan, G. G., and Allakhverdyan, M. A. (1957). Concerning the analysis of changes in light sensitivity of man's visual analyser under the action of cortical excitation and inhibition. *Problems of Higher Activity and Compensatory Adaptations*, Academy of Sciences of the Armenian S.S.R. Erevan, No. 2.

Ilyina, G. N. (1958). The manifestation of extinguishing and differential inhibition in the reaction of "opposite sign." *Typological Peculiarities of Man's Higher Nervous Activity*, Ed. B. M. Teplov. Moscow.

Mirsky, A. E. (1936). The visual cycle and protein denaturation. *Proc. Nat. Acad. Sci. N.Y.* **22**.

Nasonov, D. N., and Aleksandrov, V. Y. (1940). *Reaction of Living Matter on External Influences.* Leningrad.

Nebylitsyn, V. D. (1957). Individual differences in visual and hearing analysers with respect to strength-sensitivity relation. *Problems of Psychology* No. 4.

Pavlov, I. P. (1927). *Conditioned Reflexes.* London.

Teplov, B. M. (1956). Some questions concerning the study of general types of man's higher nervous activity. *Typological Peculiarities of Man's Higher Nervous Activity*, Ed. by B. M. Teplov. Moscow.

Ushakov, B. P. (1954). Parabiosis of the muscle and the problem of correlation between functional and substantial changes under excitation. *Achievements of Modern Biology*, **38**, 6.

BINOCULAR VISION WITH TWO STABILIZED RETINAL IMAGES

BY

R. W. DITCHBURN and R. M. PRITCHARD

From the University of Reading

Experiments are described in which each eye is presented with a target whose image remains on the same part of the retina when the eye moves. The patterns presented to each eye may be similar and may be placed on corresponding parts of the retina or may be placed in non-corresponding positions: alternatively, different targets may be presented to the two eyes. Each pattern fades intermittently. Sometimes both are seen together and sometimes both fields are dark at once. There is a small negative correlation between the times of clear vision with the two eyes. When corresponding areas of the two retinas are illuminated with red and green light respectively, the composite colour (yellow) is never perceived with steady illumination. When two similar patterns are in nearly corresponding positions there may be subjective fusion. With two different targets there is sometimes a subjective impression that the two patterns move with respect to one another even though their positions on the retina are fixed.

Introduction

In an earlier paper (Ditchburn and Pritchard, 1956) we described an apparatus for producing a visual target whose image remains on the same part of the retina even when the eye moves. Such an image is called a *stabilized retinal image*. The apparatus is shown in Figure 1(*a*) below. The subject wears a tightly fitting contact lens carrying a stalk on which is mounted an optical system which produces an interference pattern (Fig. 2(*b*)). The pattern is located at an infinite distance and an image of the pattern is formed on the retina of the unaccommodated eye. This image is fixed on the retina provided that the angle between the optic axis of the crystal and the visual axis remains unchanged. Riggs, Armington and Ratliff (1954) have investigated the possibility of movement of the contact lens with respect to the eye. They conclude that the natural movements of the eye do not significantly move the contact-lens (except possibly a few of the largest movements). This problem has also been studied by several workers in this laboratory. The results, which will be published elsewhere, support the conclusions of Riggs *et al.* It also appears probable that movements of the retinal image due to micro-fluctuations of accommodation in the emmetropic eye are too slow to affect the results of our experiments. Variations of retinal luminance due to pupil fluctuations (Hippus) are also small provided the apparatus is well adjusted. In the earlier experiments one eye viewed a stabilized image while the other was occluded. It was found that:—

(*a*) the pattern was initially seen very clearly;

(*b*) after a few seconds it faded to a grey field;

(*c*) after another short interval the field went black and remained so for up to 10 sec.;

(*d*) following a convulsive movement of the eyes the pattern was seen again very clearly. The whole cycle was repeated again and again.

The convulsive movement, which is larger than any natural movement, is probably sufficient to cause mechanical destabilization. Sometimes there is partial regeneration not associated with any unnatural convulsive movement. (We have considered

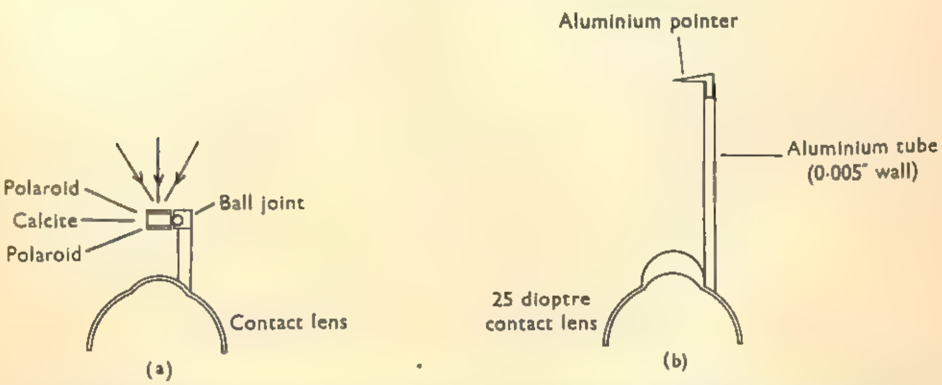
the possibility that this partial regeneration is due to a small mechanical destabilization associated with natural saccadic movements and have recorded the relation between time of flick and time of regeneration. The analysis, which is rather complicated does not indicate a simple causal relation between flicks and regeneration.) The results (*a*), (*b*), (*c*), (*d*) may be compared with other experiments on the stabilized image. (Ditchburn and Ginsborg, 1952; Riggs, Ratliff, Cornsweet and Cornsweet, 1953; Ditchburn and Fender, 1955; Cornsweet, 1956; Fender, 1956; Krauskopf, 1956; Pritchard, 1958*a* and *b*; Ditchburn, Fender and Mayne, 1959.)

It has been suggested that possibly the alternate fading and regeneration represented a form of retinal rivalry between the eye which is seeing the stabilized image and the occluded eye. If this were so then fading to a completely black field should not occur when both eyes are viewing a stabilized image. The present experiments were designed to test this point as well as to provide additional information on the possible location in the visual pathways of the effects produced by the compensation of retinal image motion.

Experimental Methods

The apparatus shown in Figure 1 (*a*) produces the pattern illustrated in Figure 2 (*b*). When white light is used to illuminate the system the rings are coloured, blue being on the inside of each ring and red on the outside. The cross is achromatic (dark grey). The angular diameters of the rings are approximately 5, 7½, 9, 10, 11 . . . degrees. None of the rings are sharp and only the first five are seen clearly. The edge of the field is grey and the diameter is 30° in the visual field. By adding a quarter-wave plate (correctly orientated) between each Polaroid and the crystal, the cross is removed, and the ring pattern obtained alone as is shown in Figure 2 (*a*). The crystal used gave a yellowish white centre.

FIGURE 1



A pattern with sharp boundaries is obtained using the apparatus shown in Figure 1 (*b*). A power of 25D is incorporated in the contact lens. An aluminium tube holds the pointer shown in Figure 2 (*d*) approximately in the focal plane of the lens, so that the subject sees an image at his far point. This is in focus for the unaccommodated eye. The angular size is as shown in the diagram.

With either of the arrangements shown in Figure 2 (*a*) or 2 (*b*) the subject can adjust the position of the centre of the pattern by rotating the crystal about the ball joint. This joint is filled with wax which is soft enough to permit movement with the hand but sufficiently stiff to hold the system in position once it has been set. Thus, if one pattern of this type is used, the subject can bring its centre to his (subjective) visual axis. If a second one is placed on the other eye, the subject can bring the two into a position which he judges to be perfect alignment or into any desired degree of misalignment.

The experimenter can adjust the "pointer" shown in Figure 2 (*d*), so that the point is seen on the subjective axis of vision. If the pointer is seen by one eye and the interference pattern by the other, it is possible to set the point so that it appears to be the centre of

the ring system, Figure 3 (*a*), or in any other desired relation to the ring system (e.g. as shown in Figure 3 (*b*)). When two ring patterns are presented both crosses may be similarly oriented or the apparatus may be adjusted so that the orientation of the cross seen by the left eye is at 45° to that seen by the right eye. When the patterns are set so that the rings appear to be superposed, the combined pattern presented to both eyes is that shown in Figure 2 (*c*).

FIGURE 2

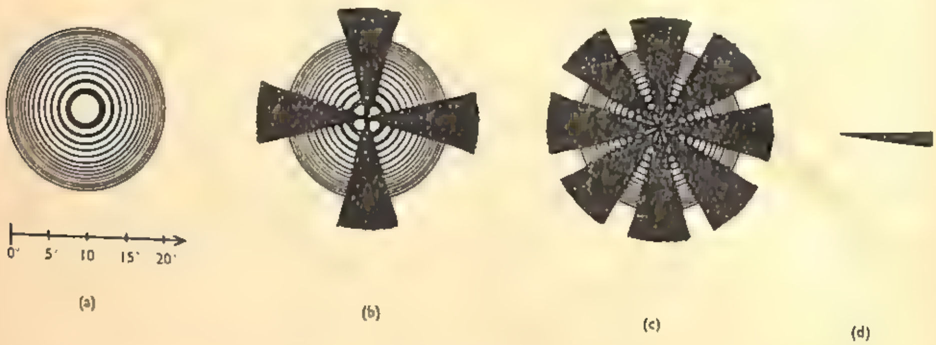
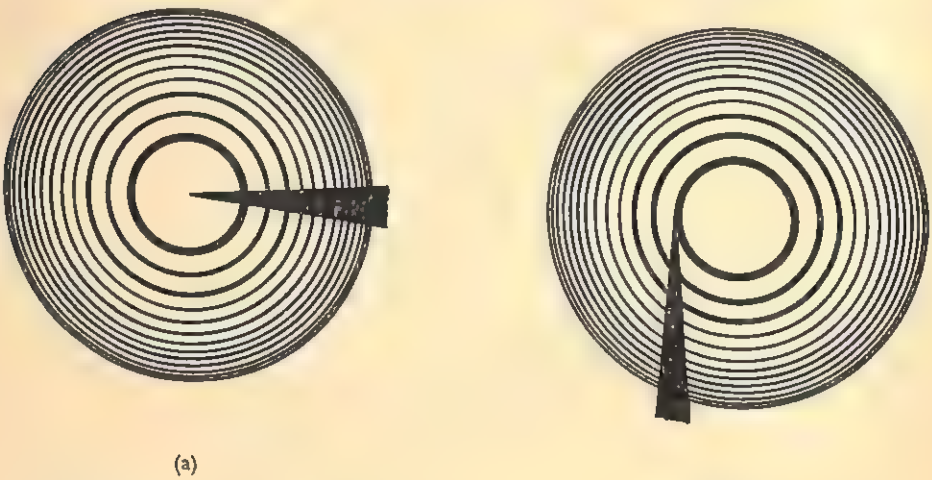


FIGURE 3



The optical systems used in the present experiments to produce stabilized images, are illuminated by magnified images of the ribbon filament of a 6 volt 108 watt D.C. lamp. The subject rests on a couch, the back of the head and shoulders being supported by a loosely fitting plaster cast. He is provided with two press-keys with which he can signal the appearance and disappearance of each image individually. These signals are registered on a three pen recorder, the third pen of which records half seconds from a standard timer.

The parameter measured in the present experiments is the Visibility Factor (V.) defined as:

$$\frac{\text{Time for which some image structure is seen}}{\text{Total viewing time}} \times 100 \text{ per cent.}$$

Results

(*a*) *Discomfort*

It has previously been noticed that subject report considerable fatigue and strain when viewing a single stabilized image for periods of about an hour in all, even though a rest is allowed every 2 min. This feeling of fatigue and strain is very much more

serious when the subject views two stabilized images. It occurs under all conditions but is worst when two patterns are misaligned by a small amount (approximately 1°). If the patterns are well aligned, or are very much misaligned and the subject accepts the situation, the arrangement is still unpleasant but tolerable. In normal vision occlusion of one or both eyes for a short period does not produce any feeling of fatigue or strain. The perceptual situation involved in viewing a stabilized image thus differs from normal vision in other respects besides the reduction of information.

(*b*) *Measurement of retinal rivalry*

The experiments described in this section were carried out with two targets aligned in the way shown in Figure 2(*c*). It is thus possible to distinguish the target presented to the left eye from that presented to the right eye. The subjects report considerable periods when the retinal images appear to fade alternately in the two eyes. However, periods do occur when both images disappear together and there are other periods when segments of both images are simultaneously visible. It thus appears that we have neither of the two extremes—complete alternation or complete independence. This is shown by analysis of Table I which summarizes data from records obtained when the subject used two keys as explained above. S_L and S_R are stabilised visibility factors for left and right eye respectively and P_M is the percentage of time for which at least one target is *not* seen. If there were complete alternation between the eyes we should have $S_L + S_R = 100$ and $P_M = 100$. P_M is always less than 100 (usually much less) and $(S_L + S_R)$ is substantially greater than 100 in 9 out of 12 experiments. We therefore conclude that, except possibly with weak blue light, we do not observe complete binocular alternation.

TABLE I

Filter		*Luminance*		*Visibility Factor*			*% time for which at least one eye is not seeing*	
Left eye	*Right eye*	*Left eye*	*Right eye*	*Left eye*	*Right eye*	$S_L + S_R$	*Measured* P_M	*Calculated* P_C
		ml.	ml.	%	%		%	%
White	White	3·5	1·0	74	79	153	44	42
"	"	"	"	80	82	162	34	34
"	"	"	"	72	83	155	44	40
Green	Green	1·0	0·1	77	74	151	48	43
"	"	0·2	0·1	64	70	134	65	55
Red	Red	"	"	56	72	128	71	40
"	"	"	"	69	65	134	66	55
Blue	Blue	"	"	47	44	91	87	79
"	"	"	"	55	55	110	73	70
"	"	"	"	57	43	100	74	64
"	"	1·0	0·5	62	75	137	57	53
"	"	"	"	74	62	136	62	54

The coloured patterns were obtained with the Chance filters OG.1; OR.1 and OB.10 and the luminances stated were measured with a flicker photometer.

If, on the other hand, the eyes were acting entirely independently the percentage of time for which both targets are seen would be $S_L S_R/100$ and the percentage of time for which at least one image is not seen would be $P_C = (100 - S_L S_R/100)$. The measured value ($P_M$) is obviously more than the calculated (P_C). This shows that

when the target presented to the left eye is not perceived the probability that the target for the right eye is also not perceived is less than it is when the target presented to the left eye is perceived. Thus there is some degree of interaction between the eyes.

The fact that there is some interaction but not a complete alternation is well established and is in accord with subjective impressions. It is, however, not justifiable to draw conclusions from the small differences of S_L and S_R for the different conditions used. The results are affected by the fact that the decision to signal that a target is not seen has a tendency to cause regeneration of one or both images. For both subjects the right eye is the dominant eye. The luminances were adjusted to make the values of S_L and S_R roughly equal.

(*c*) *Observations in the "black field"*

As stated in the Introduction, the field from time to time goes completely dark. Sometimes this happens sharply and then it always happens in both eyes simultaneously. It also ends suddenly in both eyes when the subject has the impression of a convulsive movement. In earlier experiments subjects have sometimes reported that when the target region "went black," weak unstabilized images either in the same eye or in the other eye also disappeared. It seems very probable that this general failure of visual perception is central rather than peripheral.

(*d*) *Misaligned targets*

As stated above, the subject can adjust the patterns so as to produce what he judges to be perfect coincidence of the centres. It is unlikely that his control gives an adjustment of much better than 30 min. of arc. The subject is thus unable to detect a misalignment of about 20 min. of arc.

When the targets are misaligned by about 1°, the subject can no longer fuse them. He attempts to do so and makes irregular and unco-ordinated movements of 10°. He reports the situation as painful but is unaware that such large eye movements are occurring. There are, of course, no corresponding movements of the retinal images. When the targets are misaligned by 5° the subject is more comfortable and can decide to pay attention to one pattern or the other. If one pattern is now continuously mechanically destabilized by disturbing the contact lens, the other pattern continues to fade and regenerate.

(*e*) *Apparent movement*

When the pattern of rings is seen in one eye and the pointer in the other, it is possible to set the pointer so that it just touches the innermost ring. The subject reports that the relation between pointer and ring does not remain constant over a period of a minute or so. It is possible that this effect is due to a slip of one lens on the sclera but this seems very improbable (for reasons given in the Introduction to this paper). If the eyes are moved deliberately, the relation between ring and pointer alters. This may be due to slip when the voluntary movement is very large but probably not when the voluntary movement is small.

(*f*) *Colour addition*

In normal vision it is found that, under suitable conditions, some observers can obtain additive colour mixing. When, for example, a patch of one retina is illuminated with red light and the corresponding patch of the other retina with green light, the additive colour (yellow) is perceived. We were able to obtain this effect with unstabilized images both with the ring systems and with small patches of light. The additive yellow colour was never observed with stabilized images. Sometimes one

colour was seen alone and sometimes patches of red were seen adjacent to patches of green. In general, when one pattern is illuminated with red and the other with green, the centre is perceived as green with a red surround. If one is illuminated with red and other with blue then the composite pattern has a blue centre. There are also occasions when part of the outer ring system is perceived as red and another part as green. All colours appear desaturated. If intermittent illumination (about 2 c.p.s. with the same phase in each eye) is used, the composite colour is then seen. This is in accord with other experiments which show that if stabilized images are viewed in low frequency intermittent illumination normal visual perception is restored. (Ditchburn and Fender, 1955; Cornsweet, 1956; Ditchburn, 1956; Fender, 1956; Pritchard, 1956*b*).

Discussion

When a subject views a single stabilized image there are periods when the pattern is not seen at all and periods when some part of the pattern is perceived. The latter periods are described as "periods of vision" but it should be emphasized that during these periods vision is not normal. During these times, the image appears rather faint and diffuse generally, part of the pattern may be missing or some parts of the field may appear clearer than others. Sometimes distortions of the pattern may be "seen." These observations cannot be expressed quantitatively but they do show that during the "periods of vision," the complete visual signals associated with normal vision are not being received. The fragmentary signals are being interpreted with difficulty and not altogether correctly. The observations that a considerable degree of misalignment cannot be detected (see (*c*) above) and that the additive colour is not perceived (see (*f*) above) suggest that the brain is unable to combine and compare the signals from the two eyes as it does in normal vision. These observations would readily be explained if the two patterns were not seen simultaneously but we find that, although there is some degree of alternation, there are some periods when both patterns are seen. When one eye receives green light and the other red, the additive colour is not perceived but in some regions of the visual field red is perceived and in adjacent regions green. This suggests that at any one time and place in the combined visual field the signals from one retina dominate to such an extent that signals from the other are excluded from perception. It is easy to understand that if a similar situation exists when two white patterns are presented, accurate detection of misalignment would not be possible. The above interpretation suggests that retinal rivalry may have to be considered in relation to local regions of the visual field as well as for the visual field as a whole.

This investigation was supported by a grant from the Medical Research Council and also by a research grant, No. B–1233, from the Department of Health, Education and Welfare, Public Health Service, U.S.A.

References

Cornsweet, T. N. (1956) Determination of the stimuli for involuntary drifts and saccadic eye movements. *J. Opt. Soc. Amer.*, **46**, 987–93.

Ditchburn, R. W., and Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature*, **170**, 36–7.

Ditchburn, R. W., and Fender, D. H. (1955). The stabilized retinal image. *Opt. Acta*, **2**, 128–33.

Ditchburn, R. W. (1956). Eye movements and visual perception. *Research*, **9**, 466–71.

Ditchburn, R. W., and Pritchard, R. M. (1956). Stabilized interference fringes on the retina. *Nature*, **177**, 434.

Ditchburn, R. W., Fender, D. H., and Mayne, Stella (1959). Vision with controlled movements of the retinal image. *J. Physiol.*, **145**, 98–107.

FENDER, D. H. (1956). The function of eye movements in the visual process. *Ph.D. Thesis, University of Reading.*

KRAUSKOPF, J. (1956). The effects of retinal image motion on contrast thresholds. *Army Med. Res. Lab. Rep., No. 221, Ft. Knox, Ky.*

PRITCHARD, R. M. (1958a). Visual illusions viewed as stabilized retinal images. *Quart. J. exp. Psychol.*, **10**, 77–88.

PRITCHARD, R. M. (1958b). Studies of visual perception with a stabilized retinal image. *Ph.D. Thesis, University of Reading.*

RATLIFF, F., and RIGGS, L. A. (1950). Involuntary motions of the eye during monocular fixation. *J. exp. Psychol.*, **40**, 687–701.

RIGGS, L. A., RATLIFF, F., CORNSWEET, J. C., and CORNSWEET, T. N. (1953). The disappearance of steadily fixated visual test objects. *J. Opt. Soc. Amer.*, **43**, 495–501.

RIGGS, L. A., ARMINGTON, J. C., and RATLIFF, F. (1954). Motions of the retinal image during fixation. *J. Opt. Soc. Amer.*, **44**, 315–21.

SHORTER ARTICLES AND NOTES

A SIMPLIFIED METHOD OF TREATMENT FOR SIZE-CONSTANCY DATA

BY

R. C. OLDFIELD

From the Institute of Experimental Psychology, Oxford

The traditional method of expressing the result of a size-constancy experiment is by calculation of the Thouless (1931) index

$$T = \frac{\log P - \log S}{\log R - \log S}$$

where P is the ratio of the "phenomenal" size of the two objects compared, S that of their stimulus, or perspective, sizes, and R that of their real sizes.

While this formula in itself is satisfactory, there are advantages in applying it in the manner outlined below.

(1) The introduction of P is sometimes confusing to students (and others), especially as in any actual experiment it is arranged that there is an equality match between the two apparent sizes: $P = 1$ and $\log P = 0$. It seems better to define T solely in terms of physical quantities which are actually measured, it being understood that the definition applies to the match condition. Thus we may put

$$T = \frac{\log S}{\log S - \log R} = \frac{\log S}{\log \frac{S}{R}}$$

(2) The question of whether T remains constant for different distances of the standard (and/or variable) is not often explicitly discussed. It is commonly taken to do so, within any ordinary range of distances such as those used in a laboratory experiment. Certainly in a student's laboratory exercise or in simple determinations of the index such for instance as those sometimes made in a clinical investigation, it is sound practice to make measurements for various distances of the standard. In such cases it is convenient to plot all the data on a single graph in a way that allows T to be immediately read off, while at the same time its distance-invariance may be directly confirmed, or departures detected.

The following simple procedure satisfies these requirements, and provides a useful exercise for the student in the use of graphical methods.

We may consider, for concreteness, the familiar type of experiment in which a variable diamond form at a fixed distance is compared with another, standard, one of fixed dimensions.

Let A_s be the measured dimension of the standard and A_v that of the variable which is matched with it. Let D_s and D_v be the distances of standard and variable from the subject's eye.

$$\text{Then } S = \frac{A_s/D_s}{A_v/D_v} = \frac{A_s}{A_v}\cdot\frac{D_v}{D_s}$$

$$R = A_s/A_v = y \quad \text{say}$$

$$S/R = D_v/D_s = 1/x \quad \text{say}$$

$$\text{and } T = \frac{\log y/x}{\log 1/x} = \frac{\log x - \log y}{\log x}$$

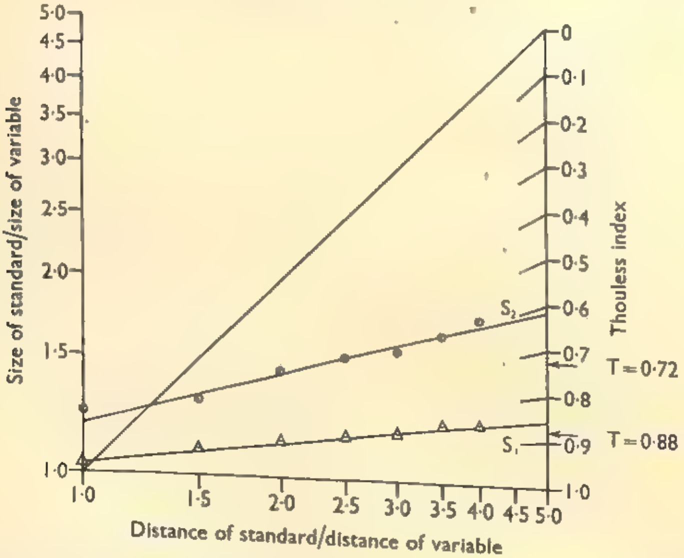
$$\text{whence } \log y = (1 - T) \log x$$

Thus if y, the ratio of the sizes of standard and variable, is plotted on double logarithmic paper against x, the ratio of their distances, the points will lie on a straight line through the origin, provided T is indeed constant for different distances of the variable. Moreover the gradient of this line is $(1 - T)$ and T is immediately determined. The X-axis represents complete constancy $(T = 1)$, and the line at 45° through the origin zero-constancy or equal-subtended-angle (perspective) judgments. Points above the latter line correspond with negative-, and below the former with hyper-constancy. Values for positions of the standard nearer to the subject than the variable lie in the left-lower quadrant.

An easy means for measuring gradients of lines through experimental points is provided by erecting an ordinate equal in height to its distance from the Y-axis and dividing it into (say) ten equal parts. The dividing points are labelled $T = 0{\cdot}9$, $0{\cdot}8$, $0{\cdot}7$ . . . $0{\cdot}1$ in order upwards from the X-axis. The value of T can then be read from the intersection of the line through the experimental points, provided the latter passes through the origin. If it does not, the parallel line passing through the origin is used.

Figure 1 shows the plot of sets of readings obtained on two subjects.

FIGURE 1



It is perhaps not uninteresting to note that the Thouless ratio, if distance-invariant, carries with it the following very simple implication. If the *apparent size* of an object be defined as the size of a variable object at a fixed distance which is judged equally large, then *apparent size varies inversely as the* $(1 - T)^{th}$ *power of*

the distance. A number of considerations, however, discourage any attempt to regard this relationship as providing a metric for visual space, the most notable being the dependence of T upon the angular separation of the objects compared (cf. Joynson, 1949).

I am indebted to Messrs. P. S. Delin and T. Morris for the data relating to Subject 2.

References

Joynson, R. B. (1949). The problem of size and distance. *Quart. J. exp. Psychol.*, **1**, 119–36.

Thouless, R. H. (1931). Phenomenal regression to the real object. *Brit. J. Psychol.*, **21**, 339–59.

THE EFFECT OF LACK OF SLEEP ON VISUAL WATCH-KEEPING

BY

R. T. WILKINSON

From the Medical Research Council Applied Psychology Research Unit, Cambridge

Sixteen subjects performed a test of visual watch-keeping twice, at the same hour of the same day of successive weeks. The test lasted 40 min. On one of these occasions the test was done after normal sleep the previous night; signals were detected at a steady rate throughout. On the other occasion the subjects had no sleep the previous night; they passed the time in innocuous leisure activities. Here a similar level of performance was maintained for the first 20 min. of the test. After this the rate of signal detection declined steadily. In the final quarter of the test the rate was roughly one third that of the first quarter.

Signals missed were divided into three categories, (*a*) signals missed while watching the display, (*b*) signals missed while not watching, and, (*c*) signals missed while asleep. Lack of sleep produced increases in all three of these categories.

These results support previous suggestions that a test has to be prolonged before performance is affected by moderate loss of sleep; they also suggest that it is the simpler, or perhaps the more predictable, tests which are more affected.

Introduction

In the past a number of studies have been made of the effects of lack of sleep on performance; as a rule periods of deprivation have ranged from 48 to 96 hr. In very few of these has loss of sleep been found to cause substantial and statistically reliable changes in performance. One reason for this may be that experimenters have tended to use relatively complex tasks, prompted, perhaps, by too literal an interpretation of Kleitman's (1939) suggestion that the main effect of lack of sleep is probably one of "fatigue of the higher levels of the cerebral cortex." Further, tests have rarely been prolonged for more than a few minutes. The result is that we have abundant evidence to show that short, complex tasks are little affected by loss of sleep even when severe.

In the present experiment a completely different type of test was used, a prolonged, simple one, that of visual watch-keeping.

Method

Sleep-deprivation

An invariable routine was observed by groups of subjects undergoing sleep-deprivation. After a normal day of duties and leisure the subjects assembled in the recreation room of the naval unit which was accommodating them. They remained there, under supervision, throughout the night, passing the time talking, writing, playing cards or table tennis. Little restriction was placed on them providing they remained awake. At 7 a.m. they performed their usual morning offices, had breakfast, and then walked about one mile to the laboratory, where they were tested at times varying from 8 a.m. to 12.30 p.m. Those tested later in the morning performed routine duties at the Naval Unit, under supervision, until due to leave for testing. Subjects who had normal sleep had exactly the same routine, except that between the hours of 11 p.m. and 7 a.m. they were free to take a normal night's sleep. The comparison is not, therefore, strictly between subjects who have had no sleep for 4 hr. and subjects deprived of sleep for 26 to 30 hr. Administrative obstacles made it impossible to supervise the non-sleepers throughout the preceding day until the 11 p.m. assembly, and so a full 26 to 30 hr. sleeplessness cannot be guaranteed. In fact, interrogation of the subjects suggested that no attempt was made to store sleep

on the preceding day. Nevertheless, the categories of sleepers and non-sleepers can more accurately be described as representing, respectively, subjects who had normal sleep and those who had no sleep during the night previous to testing.

Subjects

Subjects were naval ratings mostly between the ages of 18 and 30. All were volunteers to stay awake.

The test

The present study used one basic task which for future reference will be termed the Standard Vigilance Test. The subject sat in a cubicle and watched a ground glass screen 12 in. in diameter and 6 ft. away for the very occasional appearance of a small spot of light only slightly brighter than the back-illumination of the screen; it was approximately ⅜ in. in diameter and appeared on the screen for ½ sec. It could appear in any of eight positions equidistant on an imaginary circle of radius 4 in. from the centre of the screen. This spot of light was called the Signal; when he saw it the subject had to report the fact by pressing a morse key. The test lasted 40 min. during which the subject sat and (ideally) watched the screen continually; during this time 16 Signals appeared, irregularly spaced in time, but so arranged that four of them occurred in each quarter of the test. Spatially they occurred at random in any of the eight possible positions on the screen.

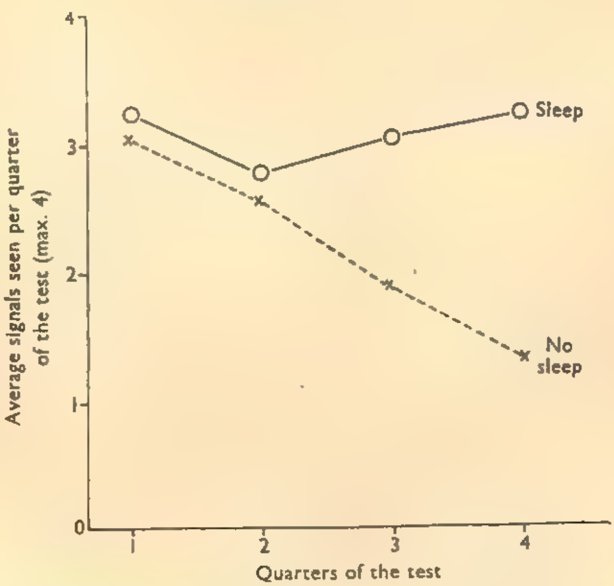
The subject wore headphones; through them he heard "white noise" at a level of about 70 to 80 db. throughout the test. This served to reinforce the sound-proofing of the cubicle in that it masked structured extraneous noise from outside. The experimenter watched the subject, without his knowledge, through a one-way window during most of the test.

On the day preceding the first day of testing the subjects were given the full instructions followed by a short run to make them familiar with the test. On the following day, before the main test, the subject was reminded of these instructions and his memory refreshed, if necessary. He was told that the test lasted 40 min., but was relieved of his watch so that he had no check on the passage of time.

Experimental conditions

In this experiment the Standard Test was given to sixteen subjects, once after sleep (henceforth S) and once after no sleep (henceforth NS). Both tests were performed at the same hour of the same weekday, one week separating the two. Two different programmes of Signals were used, one for each test. The design was arranged to balance possible

FIGURE 1



The effect of lack of sleep on number of signals seen.

differences between performance on the first and on the second tests due to "practice" effects, and also between performance on the two programmes of Signals, which could not be assumed of equal difficulty.

Results

Here we shall be concerned with the degree to which performance deteriorated in relation to normal levels as a result of loss of sleep; this will be briefly referred to as the "NS decrement." With S, subjects saw an average of 11·9 Signals out of the total of 16 presented during the test; with NS they saw an average of 8·9 Signals. This difference is significant at the 5 per cent level. (See Table I (Overall Scores).)

However, almost all of this NS decrement in total score for the whole test was due to sleepless subjects failing to maintain normal levels of performance in the later stages of the 40-min. test. (See Fig. 1.)

An index of this deterioration of performance during the course of the test is given by $Q = 2(P1 - P4) + (P2 - P3)$ where P1, P2, P3, and P4 are totals of Signals seen in the 1st, 2nd, 3rd, and 4th ten-minute periods of the test, respectively. When this index is used as a basis for an analysis of variance the NS decrement is significant at better than the 1 per cent level. (See Table I (Q scores).)

TABLE I

Analysis of Variance of Vigilance Results using Sin^{-1} Transformation of Scores of Signals seen in each Quarter of the Test

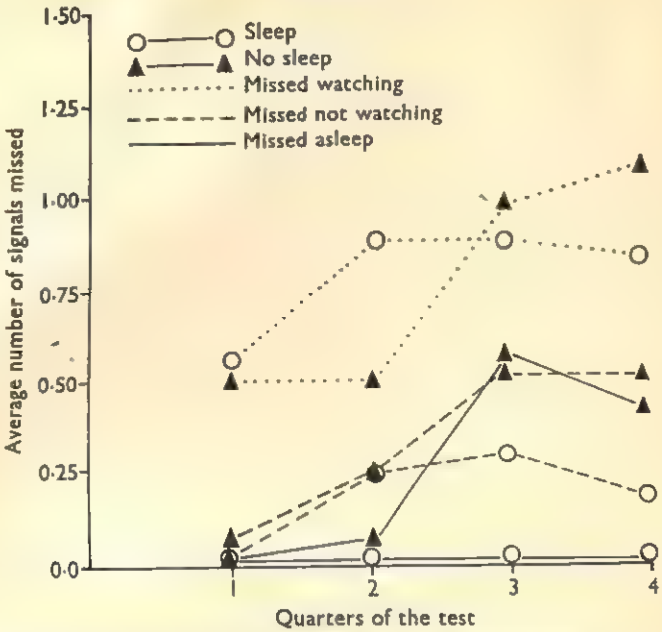
		Overall test score		*Within-test deterioration score* $Q = 2(P_1 - P_4) + (P_2 - P_3)$				
Source	*d.f.*	*Mean Sq.*	*F*	*Mean Sq.*	*F*	*Test agst.*	*P equals:* 0·05	0·01
Subjects								
Teams (T) (or S *vs* NS × D)	1	3·45	0·43	0·3	0·02	WT	4·60	8·86
Within Teams (WT)	14	8·06		19·9				
Conditions								
Sleep *vs.* No Sleep	1	33·01	6·64*	236·5	25·00**	WT × S *vs.* NS	4·60	8·86
Days (D)	1	17·26	3·47	0·1	0·00	WT × D	4·60	8·86
Residual								
WT × S *vs.* NS (or WT × D)	14	4·97		9·5				
Total	31	7·62		20·9				

But what was happening when subjects missed Signals? The experimenter recorded his estimate of the subject's attitude, as each Signal was missed, under one of three heads, (*a*) watching the screen, (*b*) not watching, with eyes shut or looking elsewhere, (*c*) not watching, asleep (eyes shut for at least 1 min. previous to the appearance of the Signal). Unfortunately these observations were taken only on the last eight subjects tested (who, nevertheless, formed a balanced group). These records have been analysed in Figure 2 to contrast the way in which subjects missed Signals in the S and in the NS Conditions.

It would appear from this that at least part of the inefficiency of the sleepless was due to their falling asleep or not looking at the screen. Nevertheless towards the end

of the test they also missed more Signals while actively watching the screen, when presumably the Signal fell on the retina but failed to evoke a response.

FIGURE 2



Analysis of how signals were missed (8 subjects only).

Discussion

When this experiment was completed and briefly reported four years ago Wilkinson (1955), it was the first instance of a large and significant effect of as little as 30 hr. loss of sleep. Since then the susceptibility of simple, prolonged tests has been confirmed by Pepler (1958) and Wilkinson (1958) for the same term of sleep deprivation, and by Williams, Lubin and Goodnow (1959) for longer periods. To the extent that it is simple, repetitive, and lacking in feedback of knowledge of results the test may be described, to use Hebb's term, as lacking in cue factor. As it is prolonged the task may become less novel and therefore be less likely, to use a mixture of the ideas of Berlyne (1950) and Broadbent (1953), to compete with other features of the stimulus totality for the subject's response.

This examination of the nature of the test suggests why it comes to be so much affected by lack of sleep. The unstimulating nature of the test and its environment may interact with the depressant effect of the stress to produce lowered level of arousal culminating, as we have seen, in overt sleep in some subjects. This trend may cause changes in the organization of the central nervous system of a kind reminiscent of Jacksonian dissolution. One result of these changes according to Jackson (1932) is that more automatic, less voluntary modes of behaviour are favoured, and this means that a task lacking in novelty and no longer exciting the subject's curiosity is less likely to be attended to. From our description of the vigilance test it would appear that it is just such a task, especially in the later stages of the test which is when we find the main impairment. To sum up then, this test may be affected so much by lack of sleep because its environment encourages lowered arousal or dissolution and its nature renders performance of it particularly susceptible to the influence of this change in central organization.

This work was prompted by a suggestion of Dr. N. H. Mackworth, it owes much to advice from him and from Mr. D. E. Broadbent. Thanks are due to the Navy for providing subjects and facilities.

References

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Brit. J. Psychol.*, **41**, 68–80.

Broadbent, D. E. (1953). Noise, paced performance and vigilance tasks. *Brit. J. Psychol.*, **44**, 295–303.

Jackson, J. Hughlings (1932). *Selected Writings of John Hughlings Jackson*, Vol. 2, 45–75. (Croonian lectures delivered at the Royal College of Physicians.)

Kleitman, K. (1939). *Sleep and Wakefulness*. Chicago.

Pepler, R. D. (1958). Warmth and lack of sleep: accuracy or activity reduced. *Medical Research Council's Appl. Psychol. Unit Report No. APU* 286/58.

Wilkinson, R. T. (1955). Effects of lack of sleep. *Interim report to Operation Efficiency Sub-Committee of the Royal Naval Personnel Research Committee, OES* 269/55.

Wilkinson, R. T. (1958). The effects of sleep loss. *Medical Research Council's Appl. Psychol. Unit Report No. APU* 323/58.

Williams, H. L., Lubin, A., and Goodnow, J. J. (1959). Impaired performance with acute sleep loss. *Psychol. Monogr.*, **73**, No. 14.

DISCRIMINATION LEARNING THEORY: EXCITATORY vs. INHIBITORY TENDENCIES IN MONKEYS

BY

GEORGE ETTLINGER

From the Psychological Laboratory, Institute of Neurology, Queen Square, London

Two methods of testing between excitatory and inhibitory tendencies in simultaneous two-choice discrimination learning by monkeys are proposed. The use of both methods yields results favouring an excitatory as opposed to an inhibitory mechanism as the unitary process in discrimination learning.

Introduction

In an important recent paper Harlow and Hicks (1957) present evidence in favour of the uniprocess conception of simultaneous two-choice learning in monkeys. They also consider what the nature of this unitary mechanism might be. "There are three possibilities: the mechanism may be excitatory, inhibitory or undefined—a mechanism X. Wisdom or cowardice would dictate mechanism X. Hebb apparently prefers excitation, and the authors prefer inhibition. At the present time there appears to be no direct test among the three mechanisms, and, this being true, choice is dictated merely in terms of which thesis best integrates existing learning data" (pp. 108–109). The present author is both wanting in wisdom because he prefers excitation and also in prudence because he ventures to propose two relatively direct methods of testing between the alternatives, excitation and inhibition.

He holds, in conformity with others, that an excitatory tendency would accrue from reward, and thus also from approach to the positive cue, inhibition from non-reward and approach to the negative cue. However, it is also suggested that if excitation underlies the process of discrimination learning then approach to the positive cue rather than non-approach to the negative cue would be the response that has been learnt and is manifested during successful activation of the formed habit; that if inhibition is basic in the process of learning then non-approach to the negative cue would constitute the final habit. Should this view be accepted then it is possible to test between excitation and inhibition by examining whether an animal approaches the positive cue or avoids the negative cue during consistently successful performance in the two-choice situation. Two ways of making this determination are proposed. In the more direct method the animal's movement leading to *correct* choice in a somato-sensory discrimination are carefully observed: the number of times tactile examination of the positive cue is followed by deliberate palpation of the negative cue prior to choice is taken as an index of the tendency to avoid the negative rather than to approach the positive cue. This principle could be adapted for visual testing. There is, however, no reason known to the author to suppose that the mechanism in two-choice discrimination learning is not the same for all modalities. In the second method a comparison is made between the increase in errors when either the positive or negative cue is spatially separated from the site of response and reward.

Method I and Results

Monkeys are most conveniently trained on a somato-sensory discrimination in a standard Wisconsin-type General Training Apparatus (Harlow and Hicks, 1957) from which all light is excluded during trials. The animal's movements can still be watched

by means of an infra-red observation device (Ettlinger and Wegener, 1958). It may then be seen that in some instances the animal's final choice is directed towards the cue it touched first without palpation of the alternative cue; on other trials it selects the original cue only after tactile examination of the second cue; or the animal may select the alternative cue after shifting from the one it touched first. As a rule discrimination training is continued until the animal makes no more than 10 errors in 100 successive trials. In such blocks of 100 relatively successful trials as at all stages of training, the animal will first chance to touch the positive rather than the negative cue in one half of all trials (on the average). Neglecting the 10 incorrect choices in such blocks, the animal will not need to shift and feel the negative cue before making the response appropriate to the positive cue and gaining the reward if it has learnt to approach the positive cue. If, however, the animal has learnt to avoid the negative cue, it might be expected to shift from the positive to the negative cue and then revert once more to the positive cue for the final response and reward.

The comparison behaviour shown by four unoperated rhesus monkeys (C1–C4) on each of two somato-sensory two-choice discrimination tests has been observed and recorded. Here reference is made only to observations taken during the final 100 trials on each test (when not more than 10 errors were made). All animals had previously been trained to the standard level of performance on at least six visual or somato-sensory discrimination tasks, but not on spatial alternation. Then they all learnt to discriminate between an hexagon and a circle in the dark; subsequently the shapes L and T served as alternative cues in a further somato-sensory test for two animals while two squares of unequal surface area served similarly for the other two monkeys. The details of the training histories of these animals and of the tests are to be published elsewhere.

When the records for these four animals are examined and all incorrect responses have been excluded from the final 100 trials on each of two tests it is found that the negative cue was touched at some time during the trial in 464 of 720 trials. Chance initial palpation of the negative cue is expected in 360 of 720 trials. The 464 examinations of the negative cue significantly exceed the chance expectation ($\chi^2 = 59{\cdot}6$, $p = 0{\cdot}001$). Nevertheless a correct choice was made in 256 trials *without* touching the negative cue. This would imply that the animals have learnt to approach the positive cue without essential reference to the negative cue. They did make deliberate reference to the negative after touching the positive cue in an estimated 104 trials. It may be hazarded that deliberate reference to the negative cue might have been made less frequently if the animals had reached a standard of discrimination more efficient than 90 per cent correct in 100 trials.

Method II and Results

Monkeys can become proficient on a two-choice simultaneous discrimination test even though the reward containers are covered by identical lids and the alternative cues are adjacent to each container. Subsequently either the positive or the negative cue can be separated by a small distance from its container. Such spatial separation of cue from response and reward decreases the efficiency of discrimination. Now if the animal has learnt to approach the positive cue then separation of the positive cue from the container should lead to greater decrement in efficiency than separation of the negative cue. If, however, it has learnt to avoid the negative cue it might be predicted (although perhaps with less confidence) that separation of the negative cue will give rise to more errors than separation of the positive cue.

A number of rhesus monkeys, including the unoperated animals C1–C4, learnt such a "spatial" discrimination as their third or fourth test. In this discrimination the

cues (three-dimensional objects) were placed on rails connecting the two reward containers. These were set 12 in. apart (centre to centre) and were covered with identical lids. Initially each cue was immediately adjacent to its container. Following completion of other testing, some of the animals reverted to this task and were retrained to the standard level of performance (not more than 10 errors in 100 successive trials). Then three new test conditions were introduced, each for only 100 trials, in a different order for different animals. *Either* (*a*) both cues were separated by 1 in. from their container (being moved 1 in. towards the other container); *or* (*b*) the positive cue alone was separated by 2 in. (again towards the other container); *or* (*c*) the negative cue was separated by 2 in. from its container. The left/right position of the positive cue and reward was, of course, always varied according to a balanced schedule. If an animal made more than 10 errors during the 100 trials of a new test condition ((*a*)–(*c*)), training was continued under the original conditions (both cues adjacent to containers) until no more than 10 errors had been made in the preceding 100 trials.

In Table I are shown the error scores for all animals tested under the new conditions. Animals C1–C4 are unoperated and served as a control group in another study (to be reported elsewhere). Following 100 trials under condition (*a*), animal C3 was not able to regain the standard level of performance within 500 trials when tested under the original conditions. Animals P2 and T1 had previously received bilateral ablations of the posterior parietal and inferior temporal regions. Three animals were tested in the light, the remainder under identical conditions in the dark. The figures in brackets represent the order in which the various test conditions were applied. In the case of two animals it was considered desirable (and it proved practicable) to increase the separation of the cues from the containers since only small differences were obtained between performance under conditions (*b*) and (*c*).

TABLE I

Effect of Separation of Cues from Site of Response and Reward

Animal	*C2*	*C3*	*P2*	*C1*	*C4*	*T1*
	Visual modality			*Somato-sensory modality*		
Separation of cues from containers:—						
(*a*) Both by 1 in.	16 (2)	37 (1)	14 (1)	2 (3)	9 (1)	3 (2)
(*b*) Positive by 2 in.	23 (3)	—	9 (2)	9 (2)	27 (2)	4 (1)
(*c*) Negative by 2 in.	7 (1)	—	2 (3)	6 (1)	3 (3)	2 (3)
Positive by 4 in.				21 (5)		10 (4)
Negative by 4 in.				0 (4)		5 (5)

The figures not in brackets represent the number of errors made in 100 trials under a given condition of cue separation. The figures in brackets indicate the order of testing under different conditions for each animal, where (1) indicates first in order, (5) indicates fifth.

The error scores can not be evaluated statistically because of the small numbers of animals in each group when differing operative status is taken into account. Only animals C1, C2 and C4 are legitimately comparable, and also contributory, since C3 was not tested except under condition (*a*). However, inspection of the Table shows that more errors are made *in every instance* when the positive than when the negative cue is separated from the container—irrespective of the order of testing. This would suggest that both in the dark and the light the animals have learnt to approach the positive rather than to avoid the negative cue.

Conclusions

The findings based on two quite different methods of comparison therefore tend to agree that approach to the positive cue is learnt as a consequence of discrimination training. As has already been suggested, this result would lead the author to infer that excitation, as opposed to inhibition, is the underlying mechanism in this kind of learning. It must, however, also be acknowledged that mechanism X of the wise, once it has been divested of its anonymity, may yet come to find support in the present findings.

I am greatly indebted to the Mental Health Research Fund and to the Research Fund of the Institute of Neurology for financial support in establishing and maintaining the Primate Research Laboratory. To Dr. J. A. V. Bates, Dr. E. A. Carmichael and Dr. E. T. O. Slater I am sincerely grateful for their help in securing special facilities for the Primate Research, and for their continued interest in this work. Professor O. L. Zangwill gave encouragement.

References

Ettlinger, G., and Wegener, J. (1958). Somaesthetic alternation, discrimination and orientation after frontal and parietal lesions in monkeys. *Quart. J. exp. Psychol.*, **10**, 177–86.

Harlow, H. F., and Hicks, L. H. (1957). Discrimination learning theory: uniprocess *vs.* duoprocess. *Psychol. Rev.*, **64**, 104–9.

VERY BRIEF DELAY OF IMMEDIATE RECALL

BY

R. CONRAD

From the Applied Psychology Research Unit, Cambridge

In a series of experiments, Brown (1955, 1958) reported a number of important suggestions concerned with the effects of interfering with immediate recall. His results which are relevant to this Note can be summarized briefly as follows:

(*a*) A very brief delay due to interpolating a single stimulus between the end of a message and the beginning of the recall of it, led to a small memory loss.

(*b*) When the interpolated material consisted of five pairs of digits, the memory loss was very much greater.

(*c*) When the interpolated material (three pairs of digits) was delayed so that some time was available for rehearsal after the message had been presented, recall improved significantly with longer delay of the interpolated material.

The above results were obtained when in general the following conditions held:

(*a*) The message to be recalled was within the immediate memory span for most of the subjects used.

(*b*) Subjects could not predict what the interpolated material was going to be.

Conrad (1958) showed that when the length of message was such that the probability of correct recall was of the order of about 0·6 to 0·7, interpolating a single digit before recall, led to a very large memory loss, even though the nature of the interpolated digit was known to the subjects in advance. The present Note briefly describes a number of experiments in which the original message was beyond the memory span for the subjects used, and the interpolated material known. These experiments complement Brown's and add a little more to what is known about the effects of delaying recall.

The general procedure was the same for all experiments. The subjects were Naval Ratings who listened to messages consisting of seven or eight digits which had been recorded on magnetic tape at a rate of 120 digits per minute. At the time appropriate in each experiment subjects repeated back verbally what they could remember of the message. This was written down by the experimenter and scored right or wrong. For a message to be scored right, the correct digits had to be reproduced in the correct order.

The digit messages were derived from tables of random numbers with the constraints that the digit nought was not used, no digit occurred twice in a message and obviously easy sequences were avoided.

Experiment I

Procedure

Sixteen subjects listened to 8-digit messages, arranged in two blocks each of ten messages. There were two conditions and although each subject heard all twenty messages in the same order, the instructions associated with the two blocks were alternated between subjects. Conditions were therefore not exclusively linked to a particular set of messages. In one condition, subjects were told to repeat back what they had heard immediately. In the other condition, they were again to repeat back what they heard, but always preceded by the digit 0. Thus the correct response to a message 74192863 would be 074192863.

Results

Without delay, the mean number of correct recalls was 72·5 per cent. With delay, the mean number was 37·5 per cent. No subject recalled more messages correctly with than without the delay.

Experiment II

Procedure

This was similar to that of Experiment I. There were three conditions, each consisting of a block of twenty 8-digit messages, and fifteen subjects different from those previously used did all conditions in a balanced order. The conditions were:

(*a*) Recall required immediately.

(*b*) Recall required after a 5-sec. interval during which the subject was urged to rehearse.

(*c*) As (*b*) but response always immediately preceded by the digit 0.

Results

Condition (*a*) 61·3 per cent. correct recalls.

Condition (*b*) 57·5 per cent. correct recalls.

Condition (*c*) 23·8 per cent. correct recalls.

No subject did better in condition (*c*) than in either (*a*) or (*b*). Comparing conditions (*a*) and (*b*), 8 subjects were better in (*a*), 6 were better in (*b*), and there was one tie. It seems reasonable to suppose that conditions (*a*) and (*b*) were not significantly different at any acceptable confidence level, but that (*c*) was in fact significantly worse.

Experiment III

Experiment I showed that very brief delay could lead to a considerable memory loss. Experiment II showed that with sequences of this length, little or no consolidation occurred when subjects were given the opportunity to rehearse. This suggested that when a rehearsal interval is provided, a delay at any point during the interval will lead to memory loss.

Procedure

Three conditions, each using a block of thirty 7-digit messages. Eighteen different subjects did all conditions in a balanced order.

Condition (*a*). Recall required after 10 sec. for rehearsal.

Condition (*b*). The subject spoke the digit 0 immediately after hearing the message, then followed 10 sec. for rehearsal before recall was required.

Condition (*c*). Recall preceded by the digit 0, required after 10 sec. for rehearsal.

Results

Condition (*a*) 70·0 per cent. correct recalls.

Condition (*b*) 44·3 per cent. correct recalls.

Condition (*c*) 49·5 per cent. correct recalls.

Only one subject did better at condition (*c*) than (*a*) (2 ties), and none did better at (*b*) than (*a*) (2 ties). Scores on conditions (*b*) and (*c*) were compared by means of the Wilcoxon matched-pairs signed-ranks test (Siegel, 1956). No evidence was found which could justify rejection of the hypothesis that the scores did not differ.

Conclusions

It seems likely that the most important difference between these experiments and those of Brown is in the length of message used. No subject could correctly recall

immediately ten successive 7- or 8-digit messages. Under these conditions—when subjects may or may not get the message right—a delay (or interference) of perhaps less than one second appears to be very damaging to the trace.

The effects of rehearsal also seem to be different according to whether the message length is within or outside the memory span. With our longer messages, rehearsal did not lead to improved recall. Nor did rehearsal appear to prevent the damaging effects of a delay before recall but after the rehearsal period. Indeed it seems immaterial whether the delay comes before or after the rehearsal.

In view of the doubt which currently prevails concerning the true cause of the forgetting that occurs in these conditions, and considering the brevity of the delay required to produce it, one feels perhaps that it is more useful to talk of an interruption rather than of a delay, since this suggests both a time duration and an interference. It is though, evident that in these experiments rehearsal has little or no effect of strengthening the trace, whereas with the shorter messages used by Brown, some consolidation did occur. In neither case could subjects know whether each rehearsal (assuming a series of formal mental repetitions) was correct. But one might imagine that with short messages, successive mental repetitions would be identical, and that the first was likely to be correct. There are two hazards with longer messages. First, repetition is more likely to be wrong, and secondly, if the subject is unaware of the error, the chance of subsequent repetitions becoming right would be small.

What seems to be of most immediate importance is to try to clarify the nature of the interruption that can destroy a just established memory trace. Clearly, as many experiments in this field have shown, interpolated tasks using material of a not very dissimilar kind from that of the message being remembered, can produce forgetting. This has led to theories both of interference and of decay, since interpolation takes time. Strictly, the experiments reported here go no further. But if the concept of interruption of the consolidation process is accepted as a postulate in forgetting, then we need to consider what kinds of event can interrupt in this way. We would also need to re-consider, without rejecting it, the status of retroactive inhibition theories of immediate forgetting. Common experience indicates the wide variety of the events, often of very short duration, which can damage a recently formed trace. It may be that the process of consolidation of a memory trace is vulnerable to any interruption, not merely to those which introduce, at the time of recall, competitive traces. The problem of defining an interrupting event remains.

I am indebted to Mrs. I. E. Skoulding who carried out all the testing.

References

Brown, J. (1955). Immediate memory. Unpublished Ph.D. Thesis in the University of Cambridge.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quart. J. exp. Psychol.*, **10**, 12–21.

Conrad, R. (1958). Accuracy of recall using keyset and telephone dial, and the effect of a prefix digit. *J. appl. Psychol.*, **42**, 285–8.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* London.

LIKE AND CROSS MODALITY RESPONSES IN NORMAL AND SUBNORMAL CHILDREN

BY

B. HERMELIN and N. O'CONNOR

From the Medical Research Council Social Psychiatry Research Unit, Institute of Psychiatry, London

Experiments which investigated response behaviour in imbeciles and normals, in a conflict situation between a general verbal instruction and direct stimuli, were carried out. It was found that combined verbal-motor responses did not give better results than either verbal or motor responses alone. Results in accordance with the general instruction were significantly better if the subnormal children responded in a different modality from the one in which the signals were given. While normals and imbeciles differed markedly in their ability to perform the tasks in like-modality trials, this difference was less apparent in cross modality responses.

Introduction

Lack of verbal fluency, and difficulties in verbal communication in the subnormal have been pointed out by Tredgold as early as 1908. More recently Luria (1956) and his colleagues in the Soviet Union have suggested that the severely mentally deficient suffer not only from a lack of development of the verbal system, but in addition from a poverty of connection between words and motor behaviour, which is held to be a crucial factor in which the mental processes of defectives of imbecile grade differ from those of normals.

According to this view language, in the normal child, not only serves as a means of communication but also as a mental tool, which develops to direct perception as well as motor behaviour. Speech, whether as instruction or self-instruction according to Luria, modifies and changes the perceptual field making designated features stand out, thus gradually enabling the child to generalize and abstract. Severely defective children on the other hand were shown experimentally to be unable to modify the perceptual field by verbal instruction. Another aspect in normal children's development is the verbal regulation of the inhibition of motor responses. Instructions to the young child tend to provoke general excitation but not to evoke appropriate responses. For example, if the young child is asked to press a rubber bulb and say "yes" to a red light, and to say "no" and not to press to a green one, he will give a pressure response regardless of whether he says "yes" or "no." His own speech acts as an excitatory impulse which can promote but not yet inhibit the response. At this stage his own speech is not regulative. Speech and motor behaviour are as yet not integrated, and the excitatory impulse from his own words sets off the motor response. At three years the normal child who accompanies direct signals by his own speech succeeds in responding discriminatingly and words assume a regulative function. This distinction between regulative function and direct excitation of motor impulses from speech is a central one in Luria's theory.

As a working hypothesis such a view has the advantage that it lends itself more readily to experimentation and remedial training, than one which regards verbal capacity as closely dependent on intelligence as a primary ability. If imbeciles are severely handicapped in using words in their directive function to guide behaviour, experiments should aim to define the conditions under which such impairment might

at least be partly overcome, and which would enable them to compensate for existing deficits. Luria reports that patients suffering from Parkinson's disease could, for instance, overcome their incapacity to continue regular pressure of a rubber bulb if they directed their own activity by counting 1, 2, 3, etc. Similarly, such patients could solve a problem of the kind "how many wheels has a car?" by tapping four times even though they were otherwise unable to tap. In this way words could activate and support an otherwise deficient motor system.

Although the impairment of imbecile behaviour seems to be general, affecting many activities, it seems to be most severe in the verbal system and in those activities which involve coding, classification and the use of symbols. Motor activity can often be substantially improved with suitable training methods. Therefore, it was thought possible that the combined verbal-motor responses of imbeciles might show improvement if compared with verbal behaviour alone. The assumption underlying this hypothesis was that a relatively intact motor system might act as support and reinforcement for verbal behaviour just as the speech system supported motor behaviour in Parkinsonism.

The present investigation was based on observations by Tikhomirova (1956) and Nepomnyashchaya (1956) concerning the rôle of speech and motor behaviour in the severely subnormal. Their studies show that imbecile children, if placed in a "conflict situation," in which responses required by verbal instructions are contradictory to those arising from the impulses set off by direct stimuli, tend to respond to the direct rather than to the verbal signals.

In our experiment the subjects were presented with verbal as well as with non-verbal stimuli. A verbal or a motor response as well as combined verbal-motor responses had to be given. The responses had to be made in accordance with a general verbal instruction, which was in contrast to the impulses received from the direct signals.

Method

Forty imbeciles and 20 normals were matched for mental age. The imbecile group was aged 9 to 15 years, mean C.A. 12 years 6 months, mental ages ranged from 5 to 7 years, Stanford Binet I.Q.s ranged from 31 to 50, mean = 41·5. Only children who had no severe motor handicap, who had some speech and who could count at least up to three were included. The normals were 5 to 7 year-olds from an L.C.C. infant school. Each child was repeatedly told what was required of him and the task was demonstrated. The child was also encouraged to repeat verbally the "rules of the game."

Each group was divided into four subgroups. In subgroups one and two the stimulus was one or two pencil taps, in subgroups three and four the stimulus was verbal counting "one" or "one, two." The responses to be made by each group in each stage of the experiment are shown in Table I. Throughout the experiment the subjects had to respond by making two responses if one signal was given and one response when two signals were given.

TABLE I

		Responses		
Groups	*Stimulus*	*Trials* 1–10	*Trials* 11–30	*Trials* 31–40
1	Tapping	Tap	Tap and Count	Tap
2		Count	Tap and Count	Count
3	Counting	Count	Count and Tap	Count
4		Tap	Count and Tap	Tap

Results

Contrary to expectation neither the modality in which the stimuli were given nor the modality in which responses were made had any consistent influence on the results. Significant differences in the ability of the imbeciles to follow a general verbal instruction and to ignore a direct stimulus were, however, evident between those groups who had to respond in the same way in which the stimulus was given, and those who used a different mode of response. If they had to respond with the opposite number of taps to those given by the experimenter, the subjects were unable to do this. Those, however, who were asked to count aloud in opposition to the number of tapping stimuli they received, obtained better results. Similarly, those subjects who were asked to tap in response to the experimenter's counting, did better than those who had to respond by counting aloud themselves. No such differences were found with the normals. Mean frequencies of correct responses for any five trials are given in Table II. As can be seen from this Table, normals respond with a similar number of correct responses under any of the four conditions.

TABLE II

Mean Number of Correct "Opposite" Responses out of a Maximum of 5

Stimulus	Trial	*1–5*	*6–10*	*11–15*	*16–20*	*21–25*	*26–30*	*31–35*	*36–40*
		Responses							
Tapping		*T*		*T + C*		*T + C*		*T*	
	Imbeciles	1·9	2·8	2·8	1·6	2·5	2·0	2·2	1·6
	Normals	4·4	5·0	4·2	4·8	4·8	4·8	4·6	5·0
		C		*T + C*		*T + C*		*C*	
	Imbeciles	3·4	2·2	3·1	2·0	1·9	2·2	3·7	2·8
	Normals	3·8	4·2	4·4	4·8	4·2	4·0	4·2	4·0
Counting		*C*		*C + T*		*C + T*		*C*	
	Imbeciles	2·2	1·3	2·8	1·4	2·2	1·5	1·4	1·6
	Normals	4·6	4·6	3·8	3·8	4·4	4·6	4·2	4·6
		T		*C + T*		*C + T*		*T*	
	Imbeciles	3·8	2·8	3·7	2·3	2·9	2·7	3·6	2·3
	Normals	4·2	3·8	2·0	4·0	4·0	3·8	3·6	3·8

Out of an optimal number of 40, the imbeciles gave 22·7 correct cross modality and 15·9 like modality responses. For the normals the corresponding scores were 31·4 and 36·1. Analysis of variance of the imbecile data showed that like and cross modality responses differ ($p = 0·001$). The difference of like and cross modality in normals was not significant. Comparing the frequency of correct opposite responses of imbeciles and normals, a highly significant difference in the like-modality trials was found, while the inferior performance of the imbeciles in the cross modality groups was less marked. There were no differences in any of the groups between the ten responses preceding and the ten following the 20 combined verbal-motor response trials, nor were the combined responses significantly different from those given in only one modality.

Instruction by the experimenter that the response should be made in "opposition" to the direct signals was repeated after each set of ten trials. If the position of correct responses in any series of ten trials is examined, it becomes evident that these occur with relatively high frequency in the imbecile groups soon after the instruction

is given. This instruction, however, gradually tends to lose its effectiveness and the difference between the number of correct "opposition responses" in the first and second half of the series of trials following instruction is statistically significant. No such difference was apparent with the normals.

Luria's observation that little connection between speech and motor behaviour is evident in imbecile subjects was confirmed in this experiment. If the experimenter gave one pencil tap the subject often gave one tap himself while saying: "one, two." When being told "to make it two taps as well as counting two" he often made one or two correct responses, and then either reverted to his previous behaviour or changed the verbal response in such a way that it reflected his motor behaviour.

It could be objected that such behaviour as was observed in the experiment was a direct function of the rather complex verbal instructions, and that a change in its verbal formulation might result in a change of behaviour of the subjects. A further experiment was therefore carried out, using the same task as previously, but instead of instructing the subject "to do the opposite of what the experimenter was doing," asking him to "play a game of three," that is, add so many pencil taps of his own to the experimenter's that together they would always make three.

Ten subjects were matched in age and I.Q. with the previous group. (Stanford Binet I.Q. 32–50, Mean I.Q. 42, C.A. 9–15, Mean C.A. 12·5). Only children who could count up to three were included.

Pencil tapping was used as stimulus as well as response. No cross modality trials were given. The experimenter explained to the subject that a "game of three" was to be played, and that the subject had to add as many taps of his own to the experimenter's as were needed to make three every time. The experimenter then demonstrated the task and performed it together with the child, guiding his hand if necessary for six trials.

Forty trials of one or two signals presented in random order were then given. The subject had to respond for the first ten trials by tapping, by tapping as well as counting aloud for the next 20 trials, and finally again by tapping alone for the last ten trials. Each correct response was rewarded.

The results proved to be very similar to those obtained in the previous experiment for the same task with a different verbal instruction. As no significant difference between the two experimental procedures was found, the results of the second study will not be presented in detail.

Discussion

The experiments described attempted to evaluate Russian findings that imbeciles were unable to take effective account of a general verbal instruction if this was in conflict with the sensory stimuli they received. A further hypothesis which was tested, stated that if the motor system could be activated to reinforce speech reactions, the weakness in the imbeciles' ability to use words as guides for behaviour might at least be partly overcome. It was expected, therefore, that combined verbal-motor responses while showing no improvement over motor behaviour might show such an improvement if compared with verbal responses alone.

The results show that the latter hypothesis, at least in so far as it concerns the rather complex experimental task, is untenable. A combination of verbal and motor responses gave no better results than either response modality alone when in the same modality as the stimulus. It seems that the use of motor movement to reinforce a speech response does not aid it, even when the motor system may be presumed to be the relatively more intact. The motor responses obtained here, lack the character of perceived voluntary movements and fail to facilitate stable joint activity of motor

and speech behaviour even temporarily. An alternative inference from the results would be that the correct method of supporting a speech function has not been developed in this experiment.

The decisive factor, which enables the subjects to comply in part with the instructions to respond in opposition to the direct signals, seems not to be the modality in which either stimulus or response is given. The tendencies to echolalia and echopraxia are more prominent if stimulus and response belong to the same modality. Cross modality reactions on the other hand seem to make a general instruction more effective, allowing greater mobility of response patterns. If the subject is required to tap when the experimenter is tapping, or to count when he is counting the compelling force of the direct signal is too strong to permit any other than a purely imitative response. The impulses received from the stimuli do not generalize so readily if the subject is asked to respond in a different and not in a similar modality.

The terms used above to describe the results are one of a number of possible alternatives. Kounin (1948), for example, has suggested that the rigidity of the subnormal is due to the tendency to segregate rather than to connect behaviour patterns. Thus, while one stimulus evokes its own response with regularity, this fails to generalize to a similar stimulus. This quality of behaviour, sometimes called "concreteness," is not unconnected with the results of this experiment. However, in this case a set to respond in a different modality partly enables the imbecile to give a different type of response, i.e. follow a general instruction and reduce the tendency to echolalia. The new set appears to relieve the stimulus response situation of its segregated character. What makes this interpretation ambiguous is that in the present experiment we are not concerned with a change in the stimulus situation but a change in manner of responses. The precise relationships of response mobility and shifts in the stimulus situation and its effect on imbecile behaviour will need further investigation.

The second conclusion which seems to follow from the results is that a general verbal instruction gradually loses its effectiveness. If the instruction to respond in opposition to the signals had been repeated with every direct signal given, one would expect a sustained response pattern.

The experiments reported suggest two conclusions which might be valuable in training the severely subnormal:

(1) Imbeciles find difficulty in freeing themselves from the compelling force of direct stimuli if they are required to respond in the same modality in which the stimulus is given. In such a situation they tend to imitate the experimenter's action.

(2) In order to have an effective function in regulating behaviour, verbal instructions probably have to be repeated every time a response is required. A generalized instruction gradually loses effectiveness. This second conclusion reinforces Russian findings reported by Luria (1957), that under such continuous reinforcement differential responses in imbeciles could be produced.

The authors are indebted to Dr. J. M. Crawford, Physician Superintendent of Botleys Park Hospital, and to Miss B. F. E. West, Headmistress of Granard Infants School, who kindly provided facilities for the research reported here.

References

Kounin, J. S. (1948). The meaning of rigidity: a reply to Heins Werner. *Psychol. Rev.*, **55**, 157–60.

Luria, A. R. (1956). *Problems of the Higher Nervous Activity of Normal and Abnormal Children*. Moscow.

Luria, A. R. (1957). The rôle of language in the formation of temporary connections. *Psychology in the Soviet Union*. B. Simon (Ed.) London.

Nepomnyashchaya, N. I. (1956). Some conditions of the derangement of the regulating rôle of speech in oligophenic children. Chapter in *Problems of the Higher Nervous Activity of Normal and Abnormal Children*. A. R. Luria (Ed.) Moscow.

Tikhomirova, O. K. (1956). Verbal regulation of movements in oligophenic children under a conflict between verbal and direct signals. Chapter in *Problems of the Higher Nervous Activity in Normal and Abnormal Children*. A. R. Luria (Ed.) Moscow.

Tredgold, A. F. (1908). *Mental Deficiency*. New York.

APPARATUS

THE REFLECTING SHUTTER PRINCIPLE AND MECHANICAL TACHISTOSCOPES

BY

J. A. DEUTSCH

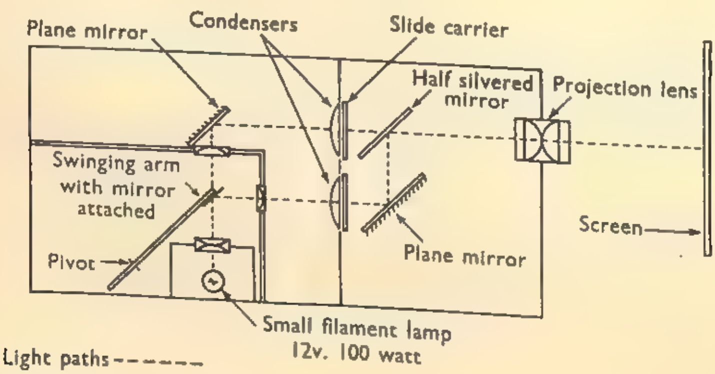
From the Institute of Experimental Psychology, University of Oxford

The principle employed in the two tachistoscopes described below is that of using a shutter which reflects light down one channel at the same time as it cuts it off from another. The principle has been applied to produce a high speed projection tachistoscope at a fraction of the cost of a comparable electronic model and a simple direct view classroom tachistoscope. The principle can be further adapted to a variety of uses.

High Speed Projection Tachistoscope

This consists of a light source, as indicated in Figure 1 (of 12 V. 100 watt, actually a car headlamp). The light from this passes through a pair of condenser lenses, forming a small image of the filament. At the point where this image is formed the light is transected by a small mirror. This passes vertically down, being attached to a pivoted arm. The surface of the mirror remains vertical as it passes downwards, and is at an angle of 45° to the light beam. Therefore the light beam takes either of two equivalent optical paths, depending on the position of the mirror shutter. When the mirror shutter is not transecting the beam and diverting it, the light passes through

FIGURE 1



another condenser lens and is then reflected by a fixed mirror at 45°, to pass through a condenser and slide-holder which can be left empty (for it forms the pre- and post-exposure field) or filled with slides to modify the steady field in any way desired. This beam then passes through a half-silvered mirror and then into a projection lens. When the reflecting shutter is in position the light beam is reflected 90° by the shutter to pass into another condenser lens and thence through another condenser and the slide to be exposed. From this the beam passes to a plane mirror set at 45° to be reflected to the other side of the half-silvered mirror and from this to the projection lens.

The model which has been built has a range of exposures from 10 to 250 millisec. This is achieved by a counterweight attached to the periphery of a wheel attached to the arm holding the mirror, where the arm is pivoted. The counterwieght is then slid to one of two positions on the periphery of this wheel. There are also two mirrors of different size which are attachable to the end of the arm to give a coarse adjustment of the range. The fine adjustment of exposure is achieved simply by releasing the dropping catch of the arm from different heights, the height being read off from a calibrated scale. This gravity drive gives a precision of timing and a constancy of operation which is difficult to attain with the type of timing circuit used in electronic tachistoscopes.

Shorter exposure times are attainable by spring-loading the pivoted arm or reducing the size of the mirror and light source. However, it has been found that a 10 millisec exposure with an equally bright pre- and post-exposure field is so low as not to be useful. Visibility at this range of speeds is a function of the brightness of the pre- and post-exposure fields.

The pivoted arm itself is kept from swinging back by coming to rest against a felt lined pad. This pad is placed almost parallel to the motion of the arm, and is forced out of parallel into the path of the motion of the arm by a light spring, so that the arm is progressively slowed down and held by the pad. However, the operator can overcome the force of the spring quite easily by pulling on the arm and reset by completing the revolution of the arm through 360°. Thus the tachistoscope can be reset without re-exposure. The slides which are used are 3 in. square. They can be photographically prepared, but the method which has been found to be most useful is to use gelatine coated glass plates (they are easily prepared) on which material is directly drawn or written in "Photopake" ink. This means that material can be prepared by the experimenter as easily as if it was drawn on paper, once there is a supply of gelatine coated glass plates.

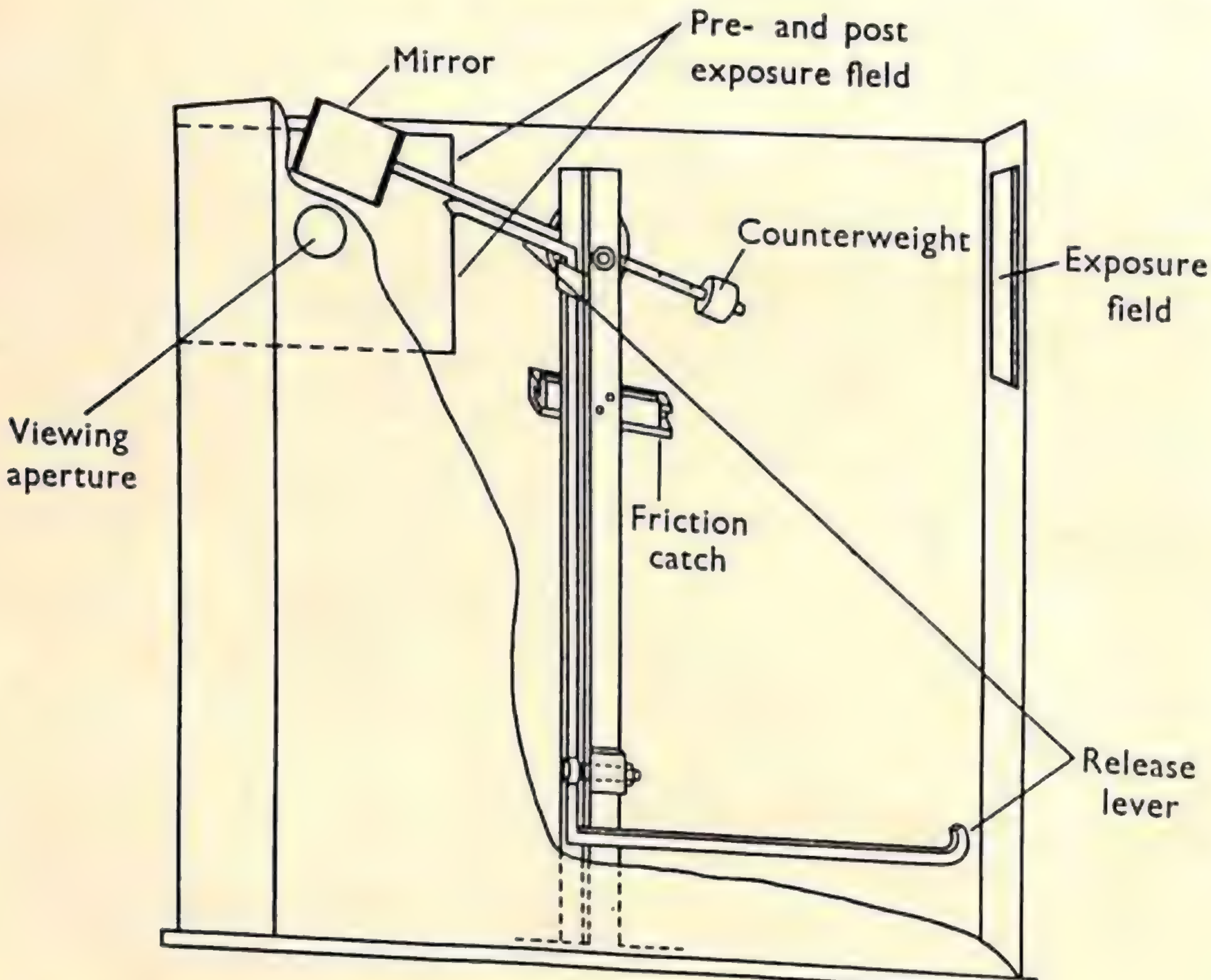
The instrument described has various advantages over the common design of electronic tachistoscopes (other than the obvious engineering ones). It has a larger and more evenly illuminated field, which can be viewed by whole audiences if need be. The spectrum of the light source employed is superior to that emitted by short decay phosphors such as magnesium tungstate, with their bluish-green colour, employed in the electronic device. In spite of the present faith in things electronic, the inherent stability of timing and illumination of the mechanical devices is much better. Where the electronic tachistoscope scores is in the versatility of its operation. It can produce double flashes with dark periods between and so on. But this is not an inherent limitation of the mechanical principle of switching the light beam instead of the light source. By the provision of a sectored wheel, motor driven if necessary such a limitation could be overcome. To produce a dark pre-exposure field, an exposure, a dark field, then another exposure followed by a dark field, a wheel would first have a dark opaque sector then a mirror, then another dark opaque sector, a gap to let the beam through and finally another opaque sector.

A Direct View Tachistoscope

In this application of the reflecting shutter principle, the pivoted arm with the mirror at the end of it (Fig. 2) moves past an aperture very close to the eye. The mirror is at 45° to the sagittal plane. The pre- and post-exposure field is set at the same optical distance as the exposure field, with a cross on it to enable fixation to take place. The two fields are easily equated for brightness. The arm projects beyond the pivot in both directions, to enable a counterweight to be slid on it to control the speed of exposure. The model made at this laboratory has a range of 40 to 120 millisec. using

a single mirror. This mirror measures 1·6 in. in the dimension which travels across the field of vision, and is 2 in. in the other dimension. Further range could be added by allowing the mirror to swing from different heights, using a more extended shaft for the counterweights and using a range of mirrors. However, for practical class use such elaboration is probably not ordinarily warranted. The drawbacks of this tachistoscope are that it is monocular (though a binocular adaptation could be designed) and that it scans the field of view. (In this last respect the projection tachistoscope is good.) However, the effect of scansion is minimized by having a fairly large mirror and that which remains seems of no great importance in classroom experiments (or, indeed, for most research purposes). One advantage which it has is that three-dimensional objects can readily be exposed. It can also be used where a fast switch from one card is desired by dropping the mirror before the aperture and keeping it there by means of a catch.

FIGURE 2



The devices described above have been found to be reliable and free from servicing difficulties, and to compare favourably with electronic devices constructed for the same purpose, for they are—what is above all important in experimenting—understood by the user.

My thanks are due to Mr. R. Shrimpton who was responsible for much of the detailed design and to whom the successful manufacture of the above instruments is due.

PROCEEDINGS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY, 1959

9th–10th April, 1959. Extended meeting at Oxford.

1st Session: "Diluted Water," by J. A. Deutsch and D. Jones. "Vision after blindness since infancy," by R. L. Gregory and J. Wallace.

2nd Session: "Recognition and recall," by R. Davis and N. S. Sutherland. "Some experiments on anagram solving," by I. M. L. Hunter. "Sequential redundancy and speed," by P. Bertelson (by invitation).

3rd Session: "Curiosity and alternation," by J. A. Deutsch and J. K. Clarkson. "The effect of local cooling on tactile discrimination," by K. A. Provins (by invitation).

4th Session: "The central control of sensory inflow," by G. Dawson (by invitation). "Neurological mechanisms in sensory integration," by J. T. Eayrs (by invitation).

13th–15th July, 1959. Extended meeting at Cambridge.

1st Session: "The effect of auditory and visual stimulation on the skin responses of schizophrenics," by P. Venables. "Problem solving," by A. R. Jonckheere. "Switching time and short-term memory in selective listening," by N. Moray (by invitation). "Two and three channel listening," by A. Taylor (by invitation). "The effect of preparatory set on immediate memory," by D. E. Broadbent.

2nd Session: "A sensitive method for the study of learning and retention," by J. Brown. "Discrepancies in measures of how much is remembered," by J. A. Deutsch. "The information-span of short-term memory," by E. R. F. W. Crossman.

3rd Session: Symposium on Primate Work. Speakers: A. Cowey, C. G. Gross, J. M. Oxbury, L. Weiskrantz.

4th Session: "Some changes in perceptual processes with advancing age," by P. A. Rabbitt (by invitation). "Degree class and attainment in scientific research," by L. Hudson (by invitation). "Some unusual handedness patterns," by O. L. Zangwill. "Experiments on simultaneous contrast," by G. C. Grindley.

5th Session: "Some electrophysiological correlates of attentive behaviour," by G. Horn (by invitation). "Quantum efficiency of rod vision," by H. B. Barlow.

30th September, 1959. One-day meeting at Birkbeck College, London.

1st Session: "The attainment of concepts," by H. Shaffer (by invitation). "Threshold changes for colours induced by anxiety-provoking stimuli," by N. F. Dixon (by invitation).

2nd Session: "Redundancy as an experimental variable," by A. Staniland (by invitation). "The effect of warning strength on the perception of a near threshold stimulus," by A. M. Halliday and R. Mingay.

5th–7th January, 1960. 12th Annual General Meeting at Birkbeck College, London.

1st Session: "Physical determinants of complementary patterns," by J. P. Wilson and D. M. Mackay. "The idea of a psychophysical scale," by C. I. Howarth. "The effect of contrast and spreading on the appearance of colour patterns and structures," by W. D. Wright (by invitation). "The Pickford-Nicolson Anomaloscope," by R. W. Pickford and R. Lakowski.

2nd Session: "Speed of response during massed and spaced learning," by J. Brown and M. Huda. "Michotte's theory of the causal impression," by T. R. Miles (by invitation).

3rd Session: "Syllabic strategies in unconstrained writing," by H. B. G. Thomas (by invitation). "Hesitation and obstruction in speech," by F. Goldman Eisler (by invitation). "Immediate memory," by E. R. F. W. Crossman.

4th Session: " 'Problem-seeking' behaviour in rats," by R. Bambridge and R. J. Audley. "A model for individual choice behaviour," by R. J. Audley and R. Pike. "Printing and the comprehension of scientists," by E. C. Poulton.

5th Session: "Cerebral stimulation and exploratory behaviour in rats," by R. Albino (by invitation). "Grooming as a displacement activity in birds," by C. H. Fraser Rowell (by invitation).

Committee, 1960: President: Professor G. C. Drew.

Editor: Professor O. L. Zangwill.

Ordinary Members of Committee: Dr. R. J. Audley (Asst. Hon. Secretary), Dr. J. Brown (Hon. Secretary), Professor K. R. L. Hall, Dr. A. M. Halliday, Dr. N. S. Sutherland, Dr. J. Szafran (Hon. Treasurer), Dr. A. M. Uttley, Dr. D. Vowles, Mr. A. Watson.

The following have been elected to membership of the Society:

J. Annett, D.Phil., Institute of Experimental Psychology, 1, South Parks Road, Oxford.

E. C. Cherry, D.Sc., Department of Electrical Engineering, Imperial College of Science and Technology, London.

R. Conrad, Ph.D., M.R.C. Applied Psychology Research Unit, 15, Chaucer Road, Cambridge.

J. Drever, M.A., Department of Psychology, The University, Edinburgh.

G. W. Granger, Ph.D., Department of Psychology, the Maudsley Hospital, Denmark Hill, London, S.E.5.

S. Griew, Ph.D., Department of Psychology, The University, Bristol.

Beate Hermelin, Ph.D., M.R.C. Social Psychiatry Unit, The Maudsley Hospital, Denmark Hill, London, S.E.5.

G. Horn, M.B., Ch.B., School of Anatomy, The University, Cambridge.

Stella Mayne, Ph.D., Department of Physics, The University, Reading.

N. Moray, B.A., Department of Psychology, The University, Hull.

S. Papert, Ph.D., Control Mechanism and Electronics Division, National Physical Laboratory, Teddington.

Hannah Steinberg, Ph.D., Department of Pharmacology, University College, London.

M. Treisman, B.A., Institute of Experimental Psychology, 1, South Parks Road, Oxford.

P. C. Wason, Ph.D., M.R.C. Industrial Psychology Research Group, University College, London.

R. H. J. Watson, Ph.D., M.R.C. Group for the Experimental Investigation of Behaviour, University College, London.

Visiting Foreign Members:

C. E. Buxton, Ph.D., c/o Institute of Education, London, W.C.1.

S. C. Ratner, Ph.D., Department of Zoology, The University, Cambridge.

BOOK REVIEWS

The Brain and Human Behaviour. Research Publications of the Association for Research in Nervous and Mental Disease. Vol. 36. 1958. London. Baillière Tindall & Cox. Pp. xi + 564. 100s.

Biological and Biochemical Bases of Behaviour. Edited by Harry F. Harlow and Clinton N. Woolsey. Madison. University of Wisconsin Press. 1958. Pp. xx + 476. 63s.

In America to-day, the elucidation of behaviour is a central issue. It is a field in which physiologists rub shoulders with psychiatrists, pharmacologists with experimental psychologists. Although the approach is sometimes labelled inter-disciplinary, the whole movement is better regarded as the convergence of a large number of disciplines upon a common focus. In view of what we stand to gain from a multi-disciplinary attack, every psychologist should welcome this invasion of his traditional preserve.

The two volumes under review testify not only to the rapidly growing concern with problems of behaviour but also to the high standard of much of the work to which it has given rise. The first brings together the proceedings of a meeting of the Association for Research in Nervous and Mental Disease held in New York at the end of 1956; the second is a collection of papers presented at a Symposium on Interdisciplinary Research organized in the University of Wisconsin. Both volumes contain material of the highest interest and importance to research workers in physiological psychology—indeed the present reviewer would regard them as indispensable to anyone who really wants to know what is going on in this rapidly developing field.

The Brain and Human Behaviour is mainly concerned with recent electrophysiological (and to a lesser extent neuropharmacological) studies which have relevance to clinical issues, and it is rather inevitable that such a volume should reflect contemporary preoccupations in neurology. Whereas fifteen years ago much space would have been devoted to the frontal lobes, pride of place is nowadays given to the temporal lobe and its subjacent structures. (Indeed seven of the twenty-one chapters deal explicitly with this topic.) Among contributions of special interest to psychologists may be mentioned Mrs. Brenda Milner's most competent study of the psychological defects produced by temporal lobe excision and the report by Dr. R. G. Bickford and others of retrograde amnesia produced by stimulation of the temporal cortex. Although so far observed in two cases only, this latter finding may prove of real theoretical significance.

A contribution of outstanding importance is the long essay by Dr. Denny-Brown and Dr. R. A. Chambers on the parietal lobe and behaviour. Here the authors present a most detailed report of their own ablation studies in the monkey and attempt some comparison with a variety of clinical syndromes. There is, moreover, a genuine attempt to analyse the data in terms of basic physiological conceptions—an attempt as stimulating as it is rare. Although not everyone will agree with Dr. Denny-Brown's contentions (which in any case are not always easy to grasp), this chapter has the makings of a classical contribution to the physiology of the brain.

Among other chapters of particular psychological interest may be mentioned the late Dr. Lashley's introductory survey of cerebral organization and behaviour, in which he returns—alas for the last time—to some of the philosophical implications of Behaviourism first raised by him in 1923; Dr. Jerome Bruner's essay on neural mechanisms in perception, a thoughtful—if not wholly successful—attempt to look at some of the higher aspects of perception in terms of a neurological model; and Dr. Lamar Roberts' review of the problem of cerebral plasticity in relation to speech.

The Wisconsin Symposium lacks the medical slant of *The Brain and Human Behaviour* and is unashamedly concerned with fundamental issues. To the present reviewer, its most encouraging feature is perhaps the extent to which systematic behaviour study has become integrated into the physiological study of the nervous system. (Indeed it is often difficult to guess whether the writer of any given chapter is a physiologist or a psychologist —a state of affairs unimaginable even ten years ago). Another encouraging feature is the revolt against undue rigidity in the design and technique of experiment and the attempt to cultivate a more adventurous outlook. In writing of primate ablation work, for example, Dr. D. R. Meyer states boldly that he cannot accept ". . . the current philosophy that all research must conform to the pattern of test, operate, retest and reconstruct."

He gives interesting examples of the gain to research of a less stereotyped approach. Dr. Hebb goes even further in arguing that the current approach to psychology in terms of "hypothesis testing" is a barrier to research. "We had much better be naïve and productive," he concludes, "than sophisticated, hypercritical and sterile." William James, one feels, would have agreed.

To the present reviewer, contributions to this Symposium of outstanding interest comprise Dr. Karl Pribram's account of neocortical function and behaviour, in which the author shows that he can handle ideas with no less skill and grace than he wields the scalpel, and Dr. Roger Sperry's account of neural plasticity and his own brilliant work on the effects of callosal section. There are also interesting and provocative chapters on neurochemical factors in the regulation of behaviour and on some neurological implications of intracranial "self-stimulation."

Taking the two volumes together, one has the reassuring impression that physiological psychology (by which is meant in this context the study of the brain as the instrument of behaviour) is very much on the move. Curiously, its antecedents lie less in modern neo-Behaviourism than in the older tradition of American "Functionalism," in the work of such men as James, Ladd, J. R. Angell, Loeb, Franz and Coghill. It betrays an approach more realistic, if less precise, than that of Hull or Skinner, and its findings can be more readily assimilated to those of other biological scientists concerned with the mechanisms of the brain. The present reviewer must also confess his unutterable relief to find that there is scarcely a mention of "learning theory" in either volume.

O. L. ZANGWILL.

Motivation: A Systematic Reinterpretation. By Dalbir Bindra. New York. Ronald Press Co. 1959. Pp. v + 361. $5.50.

Bindra has produced a useful compilation of the facts dealt with under the heading of motivation. His book is not a reinterpretation in any usual sense; it is an interpretation in the conventional language of learning theory and some Hebbian associationism. As the author says (p. 289) "The general point of view adopted in the writing of this book lies somewhere between the positions of Skinner and Hebb." Consequently, Bindra's attempt at a system fares no better than those from which it is derivative. Being however more attentive to the facts than his predecessors, in order to use the same notions, he has to make them more elastic and vague.

For instance, throughout the book, it is the rôle of learning and experience which is stressed. His position is often so general that it is hardly open to empirical disproof. For instance, his generalization I states (p. 64): "Some degree of repeated or continued exposure to a situation is a necessary condition for the development of activities that are goal-directed with respect to some feature of that situation." Mating in rats which Bindra concedes, occurs without learning, is denied the status of a goal-directed act and relegated to the status of a chain reflex. A different way out is taken with Epstein and Stellar's evidence that the taste of salt in adrenalectomized rats is unlearnt. Such animals adjust their salt intake on being allowed access to a salt solution in proportion to their deficit; and their appetite is as much in evidence during the first hour after being allowed access as afterwards. Bindra suggests that this is learnt, but fails to quote Thomson and Porter's (1953) study, discouraging to such a supposition. A part of the difficulty with Bindra's position is that he gives no hard and fast statement of how he conceives learning to take place or to specify what the various learning processes which he invokes might be. As a result, though he demands almost impossible standards of experimental stringency for those who would claim that certain motivational activities were innate, no such standards can be found when we examine his own position. As a result, almost any evidence is good enough for him to show that an activity is dependent on learning or experience, and explanations redolent of the palmy days of muscle-twitchism abound (for instance (p. 109) when he deals with Birch's ruffed female rats). Why such experiments are held to show that behaviour is not unlearnt is difficult to see and this applies to all the rather naïve deprivation experiments quoted by Hebb's supporters. It is difficult simply to deprive an animal of normal experience, because we invariably substitute abnormal experience and this will lead to the learning of interfering modes of adjustment. It is not in dispute that unlearnt behaviour patterns can be disrupted.

Another example of the one-sidedness of Bindra's critical acumen is his examination of the work of the ethologists on the subject of unlearnt behaviour. He quotes "properly controlled tests" of their experiments. An example of these is the work of Hirsch, Lindley and Tolman who "repeated" Tinbergen's work on the fear response to the hawk.

This "repetition" consisted of ignoring all the characteristics of speed, height and size of the hawk-model. Even the size of the retinal projection (assuming a chicken to go on nothing else) of the model would have been more like that of a very large aircraft at the altitude of a hawk. Of course a negative result was obtained.

Similarly, he quotes Riess (1950), but not Eibl-Eibesfeld who shows that Riess's result on deprivation of objects to carry and nest building was due to a disregard of the conditions under which nesting ordinarily takes place.

His inclusion of a chapter on "Arousal and behaviour" is praiseworthy, comprising as it does a discussion on the rôle of the reticular activating system. However, some of his interpretations of particular experiments seem puzzling. Thus, on p. 232 we find: "Tinbergen (1951, p. 29) has reported a study which suggests that the newly hatched chicks of the herring gull become more excited as judged by the frequency of pecking, on the presentation of a model of a herring gull head which has a red dot on the bill." Summing up such interpretations on the same page he states: "Thus a variety of studies can be interpreted as showing that changes in complex stimulus patterns lead to an increased level of arousal, and that the level of increase brought about by different patterns is different."

It seems a pity that the experiment of Mowrer and Jones (1943) is quoted (p. 170) as having shown that the ease of extinction of a response is a function of the effort required in making a response, after Trotter (1956) has shown that this result was an artefact due to the weighted lever employed in the apparatus in this experiment. It is perhaps less excusable that Hull's classic (1933) experiment is not correctly reported. The author (p. 181) errs in stating that the animals were given forced runs. Further the author states: "On hunger days the animals were led into the food goal-box, and on thirst days into the water goal-box." Hull used only one goal-box. The discussion of the blood chemistry of hunger takes no cognizance of the recent work of such theories as the glucostatic (Mayer) and the lipostatic (Kennedy) and the large amount of work connected with these. Of minor importance are the misprints on p. 8, 9 and 179, where "ethnologists" should read "ethologists." Of greater importance is the point that Freud is placed among those who have used the concept of "instinct" as contrasted with those who use "drive" (p. 6 and 307). Freud's notions about *Trieb* (which has been dubbed "instinct" by his translators and so misled Bindra) are wrong, but they strikingly anticipate the notions of Hull and, when we examine Freud's beliefs in the learnt nature of motivation, those of Bindra himself.

J. A. Deutsch.

Behaviour and Psychological Man. By E. C. Tolman. London. Cambridge University Press (for University of California Press). 1958. Pp. ix + 269. 17s. 6d. net.

This is a collection of nineteen of Tolman's published papers. The first is "A New Formula for Behaviourism" (1922), and the last is "Cognitive Maps in Rats and Men" (1948). These essays are, therefore, drawn from a long period of his work.

Many of the essays are not so much theoretical as metatheoretical, in that they contain mainly comments on the *kind* of theoretical structure which the author favours. In these chapters such problems as the amalgamation of behaviourism and Gestalt Theory are discussed. Empirical observations are here referred to as illustrations, rather than as evidence upon which any theoretical view depends. There are also three or four essays which may be said to be concerned with the metaphysics of behaviourism—with the status of consciousness, of emotion, and so on.

The remainder of the book is composed of two or three papers which may properly be called theoretical. In these chapters, experimental results on latent learning, vicarious trial and error, and related aspects of behaviour are reported. Most of this work is described in greater detail elsewhere. Finally, there is one paper, "The Acquisition of String Pulling by Rats," in which theoretical argument is related in some detail to a specific piece of experimental work.

The papers devoted to conceptual analysis and to general accounts of the nature of psychological theory were, for the most part, written twenty or more years ago. It is difficult now not to find them rather tedious. For, although they compare favourably with many similar works, psychologists have not usually been very successful in discussion of philosophical problems. It may be that this phase was necessary and important in the development of the subject. But it is with some relief that one finds today that such topics are more frequently set aside for analysis by those better qualified for this sort of thing. Further, it may be doubted whether any general specification *can* be given for psychological theories, apart from certain obvious logical requirements. Rather

we must apply the pragmatic test of whether they work in organizing and predicting observed behaviour. It is specially difficult to apply this test to Tolman's theories. But, in the more formal part of his theoretical papers, it is often difficult to see that the organization consists in more than the use of some rather misleading and ill-defined names. And, also in the more formal part of his work, it seems that the theory may predict anything or nothing in any actual experimental situation. The "schematic-sowbug," for instance, remains a creature of very uncertain habits.

This is the unsatisfactory aspect of this group of papers. But this collection is not quite a representative sample of Tolman's work. The book is to some extent redeemed by the less conceptual and more experimental papers which afford a better insight into Tolman's genuine theoretical thinking and into his experimental practice. It should not be forgotten that Tolman is largely responsible for the distinction between learning and performance, and for maintaining this distinction in what were, sometimes, rather adverse circumstances. Again, his insistence upon the basically goal-directed nature of behaviour is clearly expressed in some of these papers. This provides a valuable contrast with the theoretical views of some other authors, from whose systems purposive behaviour emerges more or less as a by-product. Whatever one believes the truth to be in these matters, it is still helpful at the present time to have the issues clearly stated.

In summary, this collection has little to offer to those already familiar with the main points of Tolman's work. They may well feel that this particular sample fails to do full justice to the important aspects of his work. Those who have not previously studied this part of behaviour theory will find selective reading valuable.

A. J. Watson.

Right–Left Discrimination and Finger Localization, Development and Pathology. By Arthur L. Benton. New York. Hoeber-Harper. 1959. Pp. xv + 185. $7.00.

Sooner or later somebody will want to write a book which gives a comprehensive account of the consequences of cerebral lesions in man in terms which are equally applicable to normal behaviour. Such a synthesis might well owe more to studies of brain damaged patients than to studies of normal behaviour, and psychology stands to gain more from the enterprise than clinical neurology. If therefore the title of Professor Benton's book appears narrow in compass it is worth remembering that he takes his cue, not from the theories and experiments of psychologists, but from a famous and puzzling neurological syndrome.

Professor Benton makes three main contributions to an understanding of Gerstmann's syndrome. He presents useful normative and developmental data, he uses methods which can be adapted to clarify and qualify the terms "finger agnosia" and "right–left disorientation" which have been in more or less uncritical use by neurologists since 1924, and he offers a theory that attempts to explain right–left disorientation in terms of impaired language functions.

The normative and developmental studies assembled here were based on normal adults and on normal and intellectually defective children. The adult studies make it clear that defects of finger localization and right–left discrimination on the type of tests used by Gerstmann are rare in normal adults. The studies of children go further and throw useful light on the development of these skills and their degree of association (in children) with other special abilities and with mental age. Some of the findings are of special interest. The most stimulating is perhaps the association observed between strength of hand preference and facility in right–left discrimination.

The most interesting and the most controversial section of this monograph is in the last three chapters which deal with impairment resulting from cerebral injury. Here Professor Benton discusses the neurological literature and attempts to integrate his normative findings with clinical evidence. A description of the various forms of finger agnosia and right–left disorientation and an account of other commonly associated parietal symptoms leads him to quite justifiable conclusions concerning the status of the four Gerstmann symptoms as a syndrome. There is, however, a somewhat unprofitable discussion of the "body schema."

The author's own studies of children and defectives become more obtrusive in his analysis of right–left disorientation and tend to steer his interpretation of the pathological states. The outcome is a somewhat simplified treatment of the issue of cerebral dominance and a more radical version of Head's views on the role of "symbolic formulation." For Head, "symbolic formulation and expression" were (implicitly) antecedent to language, and impairment of this function might result in disabilities (such as the Gerstmann

tetrad) which did not stem from language impairment as such. For Benton, it seems, "symbolic formulation" *is* language and "right–left disorientation of this (bilateral) type is an expression of aphasia." The evidence for this is, as he admits, by no means conclusive. This theory is to some extent the consequence of a tendency to mould the facts and fables of the clinical literature to the reassuringly precise forms yielded by normative studies. The relevance and value of these studies is undoubted, but it can be misleading to assume that correlations between different normal performances correspond to the lines of cleavage of selective intellectual disability following cerebral injury.

Finally, it is worth repeating that Professor Benton's monograph is more ambitious than his unpretentious title suggests. Positive suggestions concerning the psychological background of neurological syndromes are at present inevitably controversial. Such suggestions are, however, valuable and are too rarely uttered by psychologists.

MALCOLM PIERCY.

Sympathectomy. By P. A. G. Monro. London. Oxford University Press. 1959. Pp. xx + 290. 75s. net.

This monograph is the record of 10 years personal research into the anatomy and physiology of the peripheral autonomic nervous system. The material is well presented and there is an extensive bibliography. A valuable feature is an appendix describing in detail the author's experimental methods; in particular the techniques he has developed for plethysmography, automatic skin temperature recording and the determination of the E.S.R.

Dr. Monro displays clearly the complexity of the peripheral organization of autonomic functions and his work challenges several preconceptions. He provides evidence for the existence of many "alternative" autonomic pathways and regards the classification of "grey" and "white" rami communicantes as solely pre- or post-ganglionic as unjustified. Perhaps his most striking conclusion and one that is of great theoretical interest is that recovery of sympathetic functions in a limb following section of its normal preganglionic fibres may be a functional reorganization brought about by "*collateral sprouting* from intact preganglionic fibres."

Dr. Monro's observations are detailed and carefully made and this book will be invaluable to the anatomist and the surgical specialist. It will also be a valuable source book and at times corrective to the research psychologist who measures peripheral autonomic "indices" and treats them as closely correlated with complex central emotional states.

ALICK ELITHORN.

Dynamics of Behavior. By R. S. Woodworth. London. Methuen. 1959. Pp. x + 403. 36s.

In this short book Woodworth provides a supplement to *Experimental Psychology*, elaborated from a course of lectures presenting his own theoretical position. It is mainly devoted to theories of motivation, a discussion of some topics in perception and an account of learning. A number of other topics are dealt with at an elementary level: there are three pages on group behaviour, a paragraph on industrial relations, a page on "insight and invention," a sentence on genetics. The most marked deficiency, for an introduction to psychology, is the very inadequate treatment given to the physiological basis: a chapter on "The Control of Muscular Movement" has an inaccurate half page account of the physiology of muscle, as against three pages on "motion study," and nine on problems of aerial navigation.

In the section on motivation, the theories of Hull, Freud, Murray and McDougall are shortly reviewed. The compression and simplification necessary lead, perhaps inevitably, to some rather misleading formulations. In the discussion on Freud we are told "The Ego generates a drive referred to as 'the safety motive.' " Presumably this is the reality principle. Hull's secondary reinforcement is described as "anticipation of the primary reward." At one point the student is told "we cannot properly say that certain responses become attached to the physiological *need* for food," at another "the act that is reinforced becomes conditioned to the need." Exploration is dealt with at some length, but without mentioning the experiments on alternation.

The chapters on perception deal mainly with distance and orientation, the Gestalt principles and the constancies. The problem of shape recognition is not discussed, though references to the recent papers on this topic by British workers are included in the bibliography.

Woodworth is perhaps at his best on learning. He describes a number of recent experiments simply and clearly, and makes some effective criticisms of the neobehaviourist

account. Only Hull, Tolman and Skinner are dealt with at any length. Guthrie and Hebb are barely mentioned, recent attempts to devise theoretical models of learning are ignored and there is no treatment of memory.

These topics have been selected to illustrate the author's personal account of the dynamics of behaviour. "Dynamic" is not defined. At various points it is applied to "forces or factors" (which are also identified with "intervening variables," and can be introspectively reported on), drives, cause and effect, "the unifying dynamics of question and answer," "situation-and-goal set," and retention. It is mechanistic; in a not untypical argument the ghost of McDougall's vitalism is raised to be once more slain: McDougall believed the individual's choice of goals could not be explained mechanistically. Woodworth replies that natural selection has determined what goals we shall choose, and this is a mechanistic principle.

The main contention is that "behavior consists in active give and take between the organism and the objective environment. This inter-relationship may be called 'dealing with the environment.'" Woodworth proposes a "behavior-primacy theory" of motivation: the "direction of receptive and motor activity toward the environment is the fundamental tendency of animal and human behavior." Play and exploration are not secondarily motivated by organic needs but by the primary drive to deal with the environment. He points out that "need-primacy theory" cannot account adequately for individual adult or play interests, but that his theory does so by postulating "behavioral capacities" and a drive to actualize them.

Perception is described as a process of "decoding stimuli" which "serve as cues of objective facts." This is illustrated by distance perception. Learning is seen largely as a process of becoming acquainted with stimulus sequences, so that the animal "learns to take S_1 as a signal of the coming S_2, and to make preparations accordingly." Conditioned leg flexion to an unavoidable shock is described as a "preparation to 'take' the stimulus." Learning is produced by temporal contiguity and "the unifying dynamics of question and answer." The Law of Effect is replaced by "an undoubted characteristic of behavior which we may call the *ameliorative characteristic*."

To many psychologists there will be a curiously old-fashioned air about these arguments, for all the good points they contain. Perhaps it is inevitable that an elementary lecture course should contain no elaboration and development of theory, though much restatement. Exploration is indeed a topic of importance when considering motivation. It is perhaps as well to refute the idea that responses are made to isolated "raw stimuli" unaffected by the inter-relationships in the patterns of stimulation received, if this is still held by anyone. And description of behaviour in terms of expectation and preparation is more immediately attractive to the novice than some of the intricacies of S-R theory. But the discussion remains consistently at a verbal level; there is no attempt to analyse or explain the terms by reference to repeatable observations, experimental procedures or any logical or physiological system of structures or relationships. What is a "fundamental tendency of behavior," "objective facts," does "stimulus decoding" say anything more than "afferent neural interaction"? Explicit in Pavlov's work and implicit in much of that of the neo-behaviourists has been the desire to analyse behaviour into units which could be given physiological identification, to provide the basis for a reductionist explanation. Woodworth's failure to appreciate the nature and attraction of this programme probably accounts for the absence of any treatment of modern attempts to devise explanatory models in psychology, and the relative neglect of physiological foundations, in his introductory account.

The peculiarly transatlantic style may grate on a British ear. "A perfect square is more pregnant than any sort of near-square," "how blood chemistry can work on the brain has long been known in regard to respiration," "a social group [consists] of individuals, each of whom responds to stimuli emitted by the others and acts back upon the others," "the cues utilized by the pilot [were] fed into him verbally," "O's motivation is to up his score" are phrases that set the present reviewer's teeth on edge.

There is a useful bibliography containing a number of references to British work.

MICHEL TREISMAN.

Correction: In our review of *Psychopathology: A Source Book*, edited by C. F. Reed, R. E. Alexander and S. S. Tomkins (Harvard University Press) published in this *Journal*, Vol. XI, page 190, we omitted to state that the volume is distributed in this country by the Oxford University Press. We wish to express our apologies to the Oxford University Press for this omission.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XII 1960 Part 2

PERCEPTUAL CAPACITY OF THE ISOLATED VISUAL CORTEX IN THE CAT

BY

R. W. SPERRY, R. E. MYERS,* and A. M. SCHRIER†

From the Division of Biology, California Institute of Technology

The split-brain approach was utilized in 6 cats to test the degree of functional autonomy in the visual system. The central visual cortex of one hemisphere was isolated by the removal of most of the extravisual neocortex of the same hemisphere. Visual inflow was restricted to the isolated side in four cases by sectioning the crossed optic fibres at the chiasma and masking one eye and in two cases by section of the contralateral optic tract. The isolation produced severe deficits in visual performance, although all cases retained some ability to learn and to recall simple pattern discriminations. The results of two- and three-stage removals of the non-visual cortex, and of terminal section of the callosum, as well as histology of the lateral geniculate nucleus indicated that geniculo-striate damage was not the limiting factor. Removal of fronto-parietal cortex produced as much or more decrement in visual discrimination than did removal of temporal cortex.

Introduction

It has been found that the somatic cortex is capable of mediating high level perceptual learning and memory in the cat after its isolation by removal of parietal, occipital, and temporal cortex and section of the corpus callosum (Sperry, 1959). The aim of the following was to find out to what extent visual functions might survive similar isolation of the optic cortex. Visual impairments produced by bilateral removal of the frontal eye fields (Fulton, 1955) or the infero-temporal cortex (Chow and Orbach, 1957) suggest the probability of less functional independence in this system. On the other hand, Lashley's findings (1942, 1950) as well as the fibre relations of the striate cortex point to the possibility of a degree of functional autonomy.

The present study involved isolation of the visual cortex of only one hemisphere, the corpus callosum being sectioned and the visual inflow restricted to the same hemisphere. Earlier work has shown that cats and monkeys with midline section of the chiasma and callosum are unable to perform with one eye visual discriminations learned through the other eye (Downer, 1958, 1959; Myers, 1955, 1956; Sperry, 1958; Sperry, Stamm and Miner, 1956). This makes it possible to study the perceptual and mnemonic functions of the isolated visual cortex unilaterally, confining the extensive cortical removals to one side while the other hemisphere is kept intact to prevent paralysis and to serve general background function.

* Now at Walter Reed Army Institute of Research, Washington 12, D.C.
† Now at Department of Psychology, Brown University, Providence, R.I.

Method

Restriction of the visual inflow to one side was effected in four cases by sectioning the optic chiasma and masking one eye during training and testing following techniques developed earlier by Myers (1955). In two other cases (*Chr, Ltm*) it was effected by section of the left optic tract just posterior to the chiasma. The animals were adult cats selected as they became available after serving in other experiments.

The operated animals with vision restricted were taught to perform one of several simple pattern discriminations (see Fig. 1) which were stabilized by overtraining a minimum of 400 trials. Isolation of the central cortical area for vision was then carried out by removal of the surrounding neocortex in one or more stages and by transection of the corpus callosum, the order of the surgical steps being varied in the different cases as described below. Standard surgical precedures were followed throughout. Callosal section was effected by exposure and retraction of the hemisphere opposite the visual area to be tested to avoid damage to the latter. Special effort was made to preserve the larger branches of the middle cerebral artery running to the lateral edge of the isolated cortex.

Training and testing of visual discriminations was carried out by procedures described elsewhere (Sperry, Miner, and Myers, 1955). In brief, the cats were placed in a darkened training box (see Fig. 2) at one end of which were two swinging windows or doors each containing a transilluminated test figure. The test patterns were interchanged from right to left at random and the animals learned to push open the door holding the correct (rewarded) figure. Correct choices were rewarded with morsels of food; incorrect choices were signalled by a buzzer. Tests of generalized visual function were also applied, including visual placing responses, centring and following reactions, avoiding objects in walking, and others as described below.

The brains, following fixation and removal, were checked by gross and microscopic examination. The test hemisphere including the lateral geniculate nuclei was sectioned at 25 μ and stained with cresyl violet. The text figures indicating the cortical removals on a type feline hemisphere are based on examination of the removed brains and on maps drawn at the time of operation. On the medial surface the lesions were continued directly downward to the corpus callosum anteriorly and to the splenial fissure and pyriform lobe posteriorly.

Observations

Removal of non-visual cortex in one stage. In the first case, *Wzl*, the cortical removal was carried out in one stage to the extent indicated in Figure 3. The lateral extension into the auditory region was left intact in this and later cases to avoid damage to the underlying optic radiations. To casual examination the animal appeared to be blind on the affected (left) side during the first two weeks after the cortical removal. It seemed unable to see pieces of food held directly in front of it; it would not jump to the floor from a stool 40 cm. high; and it bumped into objects in walking about a room. Forced circling to the left was pronounced in the first two weeks, more so than in any of the subsequent cases. The tendency to veer to the left had almost disappeared at the end of five weeks. During these first five weeks discriminative performance through the left eye in the testing apparatus failed to rise above chance.

Two months after the cortical removal, *Wzl* had regained the ability to discriminate the light vs. dark patterns L–D. It then relearned discriminations H–V and X–O with no evidence of any retention from the preoperative training. The performance on X–O was erratic, but on H–V *Wzl* attained a consistent performance level of 92 per cent. correct responses.

Section of the corpus callosum four months after the cortical removal again abolished all discriminations up to five weeks after the operation. Beginning with the sixth week after callosal section, discriminations L–D and H–V were relearned a little easier than initially, suggesting some saving. *Wzl* learned further to discriminate a horizontal bar from an upright bar that was tilted from the vertical by stages up to 30°. Beyond this angle the discrimination broke down.

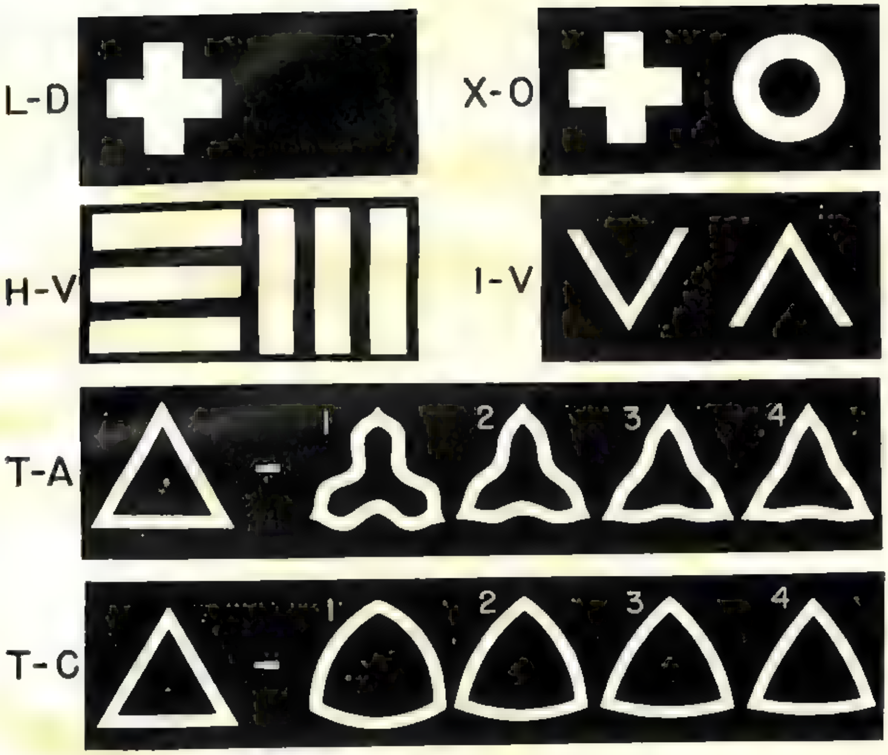


FIGURE 1

Pairs of patterns used in testing visual discrimination. In T–A and T–C only one of the distorted triangle series was paired with the equilateral triangle at left.

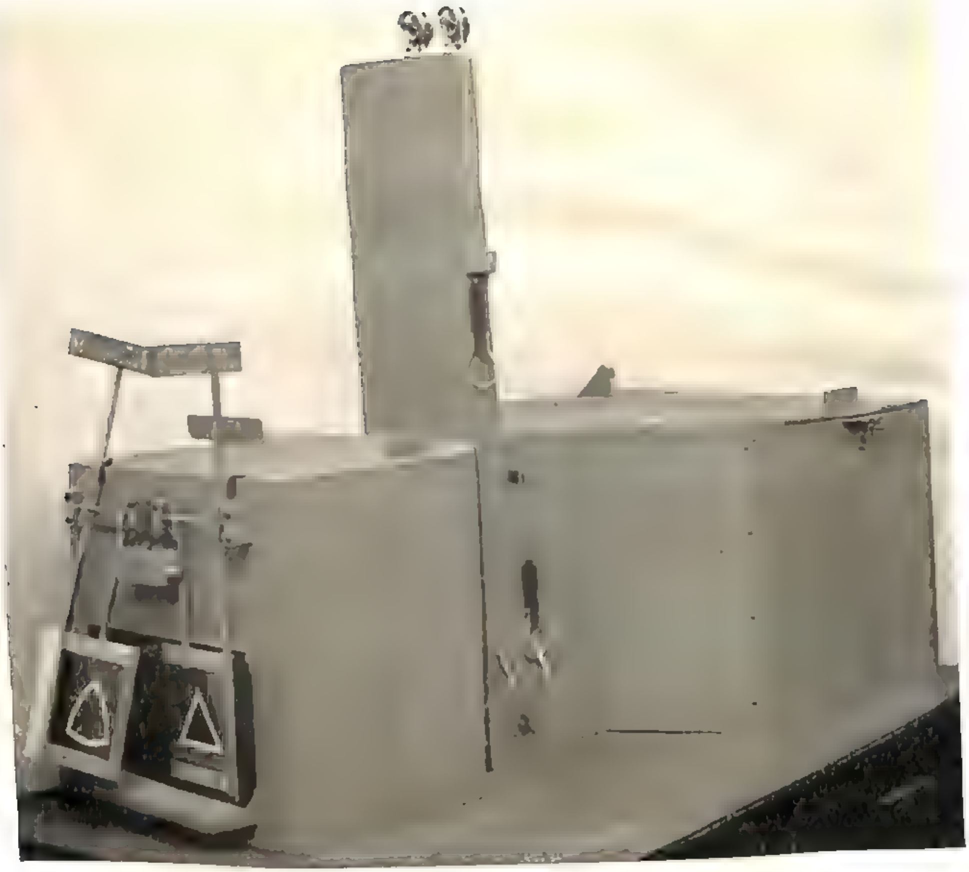


FIGURE 2

Discrimination apparatus. Starting chamber is separated from choice chamber by two sliding doors, one opaque, one transparent.

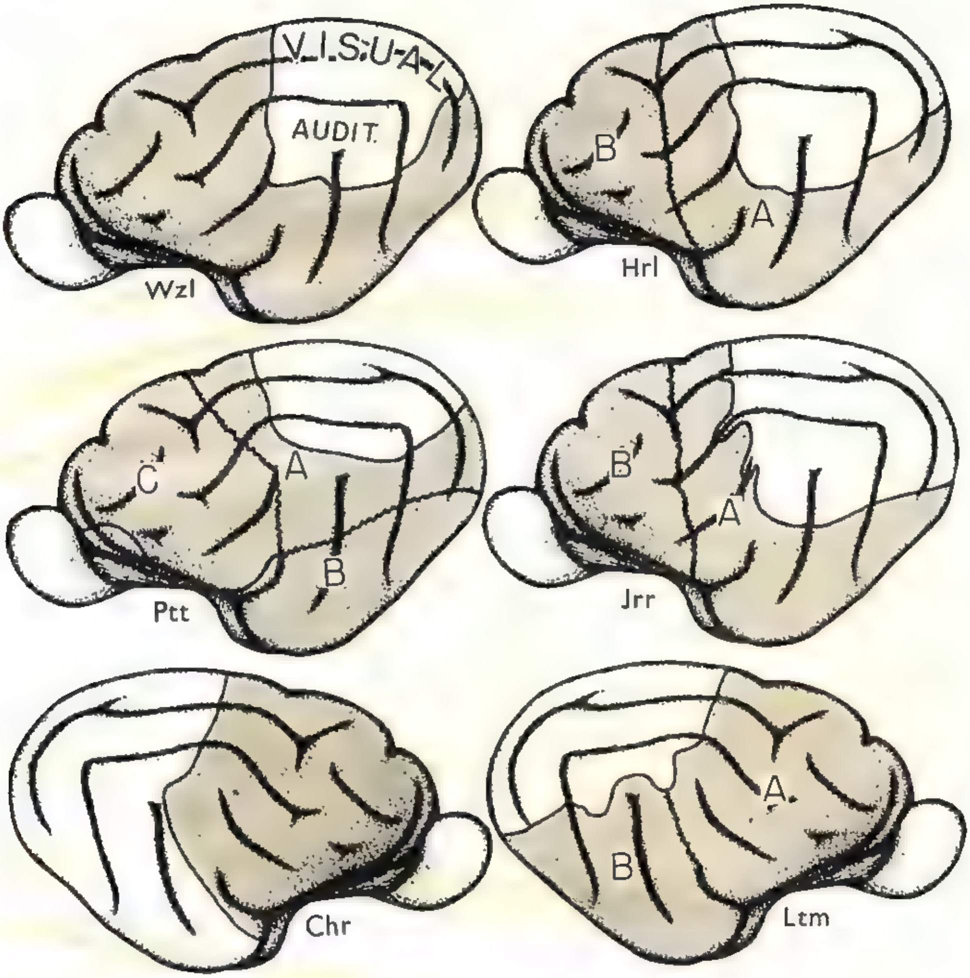


FIGURE 3

Extent of cortical removals (stippled) in test hemisphere. Where the removals were carried out in stages, the order is indicated alphabetically.

In further visual tests with the left eye conducted at three months after section of the callosum we were unable to evoke centring, fixating, or following movements of head and eyes. In walking about an unfamiliar room, the animal bumped into objects directly centred in its path. Visual placing responses were absent as were also the reaching for and striking at small objects. The animal would not jump from one table to another of equal height at 25 cm. distant and balked at jumping to the floor from a low stool. The cat seemed unable to see pieces of meat dangled close in front of its nose. In the discrimination apparatus its performance was very slow and uncertain, and it seemed to tire easily. It would typically run a good series of 9 to 15 perfect runs, then relapse with many errors.

Removal of non-visual cortex in two stages. In the next case, *Hrl*, a wide band of the left cortex bordering the area to be isolated (A in Fig. 3) was removed first so that any damage to the underlying fibre systems would be limited mostly, or entirely to this first stage operation. The preoperatively trained H–V survived this removal with only a transient drop in performance level that was recovered by the 14th day after operation.

On the 18th day following the cortical removal, the corpus callosum was sectioned. During the next ten days the animal appeared blind in the left eye. By the 20th day it was again able to discriminate H–V, and was performing 98 to 100 per cent. correct during the three weeks preceding the final cortical removal. Its visual performance in general remained relatively good.

The remaining frontal region of the left cortex was removed four months after the callosum had been sectioned. This produced an additional impairment of vision, evident both in general tests and in the discrimination box. Beginning 43 days after this operation *Hrl* was again able to discriminate H–V, and like *Wzl*, it learned to distinguish a horizontal bar from one tilted up to 30° from vertical. Other tests of general visual function administered during the 12th week after the final operation yielded much the same results as obtained in the final stage in the preceding case.

Removal of non-visual cortex in three stages. The third case, *Ptt*, was trained to discriminate H–V, X–O, and I–V following transection of the optic chiasm, and the callosal and hippocampal commissures. The first cortical removal involved the area A shown in Figure 3. Discrimination H–V survived with moderately good performance, but X–O and I–V had to be relearned. However, the relearning went rapidly, suggesting partial retention, particularly for X–O.

After consistent performance had been restored on all three discriminations, cortical area B (Fig. 3) was removed 32 days after the removal of A. Discrimination H–V and X–O survived this second removal, but I–V was lost. With retraining, the performance levelled off on H–V at 88 per cent., and on X–O at only 68 per cent. correct.

When the isolation was completed by removal of the frontal region 43 days after removal of B, the effect was much the same as the final stage in the two preceding cases. In the course of six weeks it was possible to retrain H–V and X–O, but not I–V on which there was no sign of improvement after 750 trials. In this final condition *Ptt*, like *Wzl* and *Hrl*, failed to respond visually to food dangled in front of its nose; it walked directly into objects; and in general exhibited much the same picture of visual deficiency as had cases *Wzl* and *Hrl*.

Less drastic removal, two stages. In this fourth case, *Jrr*, the cortical removals were much like those in *Hrl* except that a larger cortical area was left intact (Fig. 3). Particularly the depths of the anterior ectosylvian and suprasylvian fissures were intentionally preserved to avoid damage to possibly important fibres coursing from

the intact cortical area forward into the internal capsule and basal ganglia. In this case the callosum and anterior and hippocampal commissures, as well as the optic chiasm had previously been sectioned and the cat had been trained to perform all the discriminations shown in Figure 1.

The first step in cortical removal involved the area A shown in Figure 3. All the discriminations survived this except I–V, which was quickly restored to the criterion level of 17 out of 20 correct by the 60th trial. Discriminations I–V and T–A3 levelled off at 85 per cent. correct, the others above 90 per cent.

Two months after the removal of A, the remaining frontal neocortex was removed on the same side. Discriminations H–V and X–O were both retained although the performance was much more erratic than before the operation. T–C1 was retained and it was possible to retrain T–C2 but not T–C3. T–A3 was lost but it was possible to retrain T–A2. The performance levelled off between 79 per cent. correct on T–C2 up to 84 per cent. for H–V. The I–V discrimination was lost and could not be retrained.

In the general visual tests this case was somewhat superior to the others. It could centre its gaze on a piece of meat approximately 3 cm. in length held 20 cm. away and could follow the movement of small objects across the middle of the visual field. Visual placing reactions with the left foot were present, though they usually fell short. Unlike the others, this cat seldom bumped into objects in walking an unfamiliar room except when they were approached from the blind half of the visual field.

Visual area isolation with tract section. The experiments were repeated in two additional cases in which the visual inflow was lateralized by section of the left optic tract just posterior to the chiasma. Tract section has the advantage of leaving the crossed as well as the uncrossed optic inflow to the test hemisphere. It also obviates masking one eye during training and testing. In general the visual impairment under these conditions was of much the same order as in the preceding with the following worth special mention.

In both cases the frontal area, labelled A in *Ltm* in Figure 3, was removed first. This produced an impairment as great as or more severe than when the temporal and parietal removals were first. By the third month after this removal the discriminations H–V, X–O, and T–A1 had been recovered to a level of 90 per cent. or better.

The temporal area (B in Fig. 3) was removed in *Ltm* six months after the frontal removal. Beginning eight weeks later *Ltm* again began to perform H–V and X–O and attained levels of 95 and 90 per cent. correct, respectively, by the tenth and eleventh weeks. When the callosum was sectioned in *Ltm* six months after the temporal removal, discrimination was reduced to a level of pure chance for three weeks. Beginning on the 4th week, H–V came back and reached a level of 84 per cent. correct. Three months later *Ltm* was still unable to perform X–O and in the general tests also it displayed a degree of visual defect much like the final stage in *Wzl*, *Hrl* and *Ptt*.

In *Chr* section of the callosum five months after the frontal removal abolished pattern discrimination for three weeks, after which X–O and T–A1 were recovered. However, the performance of these did not rise above 80 per cent. correct during the next three months. The animal continued to perform so slowly and poorly that the temporal removal was not carried out.

Weak centering and following reactions, as well as visual placing, survived in both *Chr* and *Ltm*. Curiously, the callosal section in *Chr* left a pronounced asymmetry in pupillary size, the right pupil being dilated. This was still present, though less

marked, as much as a year later. The pupillary effect was absent or at least much less striking in *Ltm*. All the chiasma-sectioned cases exhibited the lasting bilateral dilation that typically follows chiasma section in the cat.

When the discrimination tests were completed, *Ltm* and *Chr* were used further in a conditioned reflex study (Voneida and Sperry, unpublished). With a flashing light serving as the signal stimulus which was paired with an electric shock to the right paw, it was possible to establish in both cases a conditioned lifting of the right paw to the visual signal.

Anatomical Checks

Examination of the removed brains showed the following: Section of the crossed fibres in the chiasma was complete in *Wzl*, *Hrl*, *Ptt*, and *Jrr*. The left optic tract was also completely sectioned in *Chr* and *Ltm*. Division of the corpus callosum and adjacent psalterium was complete with the following two exceptions: A fine strand of fibres no more than 1·0 mm. in diameter was intact at the posterior edge of the callosum in *Ptt*. The anterior fourth of the callosum was intact in *Hrl*, but it was hardened and presumably degenerate owing to removal of all its cortical connections on the left side.

Examination of right and left lateral geniculate bodies in the chiasma-sectioned cases revealed patches of retrograde degeneration on the isolated side. These degenerate areas were located on the edges, however, and in the cases most severely affected, *Wzl* and *Hrl*, the degenerate areas did not total more than half the volume of the principal dorsal nucleus. Extensive retrograde changes were apparent in the thalamus on the isolated side. The hippocampus, and amygdaloid nuclei were undamaged for the most part as was also the pyriform lobe except for minor invasion of the lateral pyriform area in three cases. On the medial surface most of gyrus proreus was intact in *Jrr*. Smaller remnants of the cingulate and proreus gyri were found on the medial surface in other cases, but no correlation between these and the experimental results was evident.

Discussion

The experiment was undertaken initially on the assumption that directionality in vision is perceived with reference to body posture, especially that of the head and eyes, and hence if all somesthetic cues for posture could be prevented from integrating with the retinal cues, it should have profound effects on visual function. Although the results are in line with this reasoning, they are by no means conclusive.

Removal of the temporal cortex produced no greater, and perhaps somewhat less impairment of vision than did removal of the frontal region. It would appear that the inferotemporal cortex does not play the critical rôle for visual perception in the cat that it does in the monkey (Chow and Orbach, 1957). The nature of the extravisual contribution is quite unknown. The cats acted as if the impairment involved the perceptual process itself rather than its motor expression.

It should be emphasized that survival of the training-box discrimination of the simpler patterns employed in this study is indicative of only a very low level of perception. Even with the entire visual cortex removed bilaterally, the cat may still respond to differences in light intensity and to moving vertical stripes (Smith, Kappauf, and Bojar, 1940). Thus, by using head or eye movements, the cats could conceivably have performed L–D, H–V, and X–O with little use of the isolated area. Actually, head movement did not appear to be used to facilitate the discrimination in this way. Eye movements were not observable.

The visual impairments were not simply a reflection of geniculostriate damage judging from the histology of the lateral geniculate body. The fact that additional impairments were produced by the removal of the distant frontal and temporal lobes after the cortex between these and the preserved occipital remnant had already been removed at an earlier date, gave further indication that the deficits were not attributable to damage either to underlying projection fibres or to circulatory or other secondary damage within the isolated cortex.

The visual deficits were least severe in *Jrr*. The intact cortex that set off this case from the others seemed to be primarily the depths of the anterior ectosylvian and anterior sylvian fissures. This points to the possibility that the cortex here or the fibre systems underlying it have special importance for vision. However, much more analysis will be needed to determine the relative importance of the various cortical areas involved. In general the results point to the probability of several or multiple extravisual factors with a "mass action" type of effect.

In a related experiment in progress somewhat similar isolation of the occipital cortex in a split-brain monkey with sparing of the geniculostriate tract has been observed (Sperry, unpublished) to produce even a greater blinding effect than in the cat.

The transient blindness of one to several weeks' duration produced on the isolated side by section of the corpus callosum was in part presumably a manifestation of diaschisis potentiated by the preceding cortical removals. Diaschisis from major neocortical damage seems in our experience not to be fully recovered before about six to eight weeks in the cat. This fits with the observed period for recovery and enhancement of evoked cortical potentials in areas affected by callosal degeneration (Franklin and Desmedt, 1957). Some of the added impairment produced by callosal section persisted well beyond this period in *Ltm* and *Chr*. This would suggest that sensory or perceptual data were being communicated over the callosum for efferent processing by the more efficient, intact hemisphere.

The visual placing response survived isolation of the visual cortex and the paw preferred was the one contralateral to the remaining somatic area despite the surgical separation of the latter from the receiving visual cortex. This fits with the observation (Schrier and Sperry, 1959) that the split-brain cat can use either paw with equal correctness in learning and performing monocular discriminations.

The work was supported by the F. P. Hixon Fund of the California Institute of Technology and by research grants from the National Science Foundation and the Southern California Society for Mental Hygiene. The valuable assistance of Lois MacBird and Alice McCulloch in the training and testing is gratefully acknowledged.

References

Chow, K. L., and Orbach, J. (1957). Performance of visual discriminations presented tachistoscopically in monkeys with temporal neocortical ablations. *J. comp. physiol. Psychol.*, **50**, 636–40.

Downer, J. L. (1958). Role of corpus callosum in transfer of training in *Macaca mulatta*. *Fed. Proc.*, **17**, 37.

Downer, J. L. (1959). Relation of optic pathways and corpus callosum to visual tracking behaviour in monkey. *Fed. Proc.*, **18**, 38.

Franklin, L., and Desmedt, J. E. (1957). Sensibilisation progressive d'une aire corticale parteillement denervée à ses afferences residuelles. *Compt. rend.*, **151**, 2204–8.

Fulton, J. F. (1955). Cerebral cortex and motor functions. In *Textbook of Physiology*, Philadelphia.

Lashley, K. S. (1942). The mechanism of vision: XVII. Autonomy of the visual cortex. *J. genet. Psychol.*, **60**, 197–221.

Lashley, K. S. (1950). In search of the engram. *Symp. Soc. exp. Biol.*, **4**, 454–82.

MYERS, RONALD E. (1955). Interocular transfer of pattern discrimination in cats following section of crossed optic fibers. *J. comp. physiol. Psychol.*, **48**, 470–3.
MYERS, RONALD E. (1956). Function of corpus callosum in interocular transfer. *Brain*, **79**, 358–63.
SCHRIER, A. M., and SPERRY, R. W. (1959). Visual-motor integration in split-brain cats. *Sci.*, **129**, 1275–6.
SMITH, K. U., KAPPAUF, W. E., and BOJAR, S. (1940). The functions of the visual cortex in optic nystagmus at different velocities of movement in the visual field. *J. gen. Physiol.*, **22**, 341–57.
SPERRY, R. W. (1959). Preservation of high-order function in isolated somatic cortex in callosum-sectioned cat. *J. Neurophysiol.*, **22**, 78–87.
SPERRY, R. W. (1958). Corpus callosum and interhemispheric transfer in the monkey (*Macaca mulatta*). *Anat. Rec.*, **131**. 297.
SPERRY, R. W., MINER, NANCY, and MYERS, R. E. (1955). Visual pattern perception following subpial slicing and tantalum wire implantations in the visual cortex. *J. comp. physiol. Psychol.*, **48**, 50–8.
SPERRY, R. W., STAMM, J. S., and MINER, NANCY (1956). Relearning tests for interocular transfer following division of optic chiasma and corpus callosum in cats. *J. comp. physiol. Psychol.*, **49**, 529–33.

GROUPING OF STIMULI AND APPARENT EXCEPTIONS TO THE PSYCHOLOGICAL REFRACTORY PERIOD

BY

A. M. HALLIDAY, M. KERR and A. ELITHORN

From the Neurological Research Unit of the Medical Research Council, National Hospital, Queen Square, London, W.C.1

When a subject is called upon to respond independently to two stimuli, the second of the two responses is often delayed if the stimuli follow closely on one another, and this has led to the suggestion that in making decisions the human operator accepts and organizes the available input information intermittently in the manner of a discontinuous servo. According to this view two nearly simultaneous stimuli can only be dealt with equally fast if they are grouped into a single decision to respond to both stimuli; otherwise one will have to wait for the attention of the central mechanism until the other has been dealt with. In the present experiment it is shown that delays in the second response are not necessary or invariable, and that the pattern and timing of the second responses when they are performed without delay differ in important respects from those to be expected of grouped responses. It is concluded that the central mechanisms concerned in the response do not possess the limitations that the single channel theory would suggest.

Introduction

It is widely held that one cannot attend to more than one thing at once, but this statement is vague in that it does not make clear what is meant by "one thing." Even simple sensations involve a complex input from many sensory receptors. However, it has been suggested (e.g. Hick and Welford, 1956) that at least the "decision process," intervening between the central arrival of a sensory stimulus and the central initiation of a response, is a process which takes place in, and has to be routed through, a single central channel, and that this mechanism acts in an intermittent manner, dealing with only one set of data during the finite time it takes to reach its decision and initiate the appropriate response. Input information is gated into the central channel intermittently and is allowed in only *before* the commencement of the period ("central organizing time") during which it is translated into an appropriate output; information arriving later cannot be used to influence the response, but must be discarded or held in store until the channel is cleared. Similarly, information relative to a decision about some other response will also have to wait until the first response is organized and initiated; only then can the organization of the second response begin. Thus even if the two responses involve quite independent sensory and motor pathways, the second will be delayed if it arrives centrally while the first is still engaging the hypothetical single central channel.

This model was evolved by a number of earlier workers (Craik, 1947, 1948; Vince, 1948; Hick, 1948) to explain the experimental data on which the concept of the "psychological refractory period" is based. Welford (1952) gave the theory a quantitative form, and considered in detail the possible further hypotheses and refinements that were needed to enable it to allow for the experimental data then available. Davis (1956, 1957) rejects some of the details of Welford's model, but accepts the main concepts outlined above and has provided some further experimental support.

In Welford's model, the lengthening of the second response, which occurs if the appropriate stimulus arrives before the hypothetical "decision process" to the first is over, depends quite simply on the difference between the interstimulus interval (I) and the length of the first reaction time (RT_1), so that the actual $RT_2 = RT_N + RT_1 - I$, where RT_N is what RT_2 would have been were there no interfering first stimulus. When plotted this relationship will be represented by a straight line with a negative slope of 45°, falling from a maximum when the second stimulus (S_2) occurs at the earliest distinguishable interval after the first. It may be added that if, rather than taking RT_1 itself, one reduces this value to allow for motor conduction time, the slope of the predicted RT_2 remains the same, though its height and length would be altered by the corresponding amount.

Not all the experimental results have fitted this prediction. Elithorn and Lawrence (1955) showed that in one situation the longest second reaction times might occur when the second stimulus came, not immediately, but 50 to 250 millisec. after the first. Moreover, in their subjects this slowing of the early second responses tended to be selectively "learnt out" during the course of the experiment.

But this is not decisive evidence against the intermittency hypothesis. Welford, in his original paper (1952), had already considered certain apparent exceptions to the "psychological refractory period" noticed by Vince (1948); he concluded, with her, that they were not true exceptions at all, but examples of "grouping," the two stimuli being responded to, not separately and consecutively, but as a single situation, i.e. "together." It is suggested that in certain circumstances this can occur when the two stimuli are separated by an interval of anything up to 50 to 100 millisec., and that it is associated with conscious or unconscious waiting for a brief period, during which the subject decides whether a second stimulus is coming almost at once.

There appear to be several possible criteria which can be adopted to distinguish grouped from ungrouped responses in this type of experimental situation. As the decision to respond to "both" stimuli is supposed to involve a single decision process, the grouped responses should be made together provided that there is no additional source of delay. Simultaneity of response is, therefore, one criterion. This simultaneity need not be absolute, since differences in conduction time in the two peripheral motor pathways will still appear. But similar differences should be found in the control mean reaction times of the two responses performed separately.

Another criterion would be the preponderance of long RT_1's associated with short RT_2's at the early intervals. This includes as "grouped" not only simultaneous responses but those responses where no "refractoriness" has occurred, and where there is evidence of waiting or straying of the attention at the time of S_1. Davis (1956) employed this criterion, calculating the product moment correlation coefficient between the reaction time of each of his pairs of responses.

The method adopted in the present paper is to plot the pairs of responses in the form of a scattergram (see Figs. 3 to 5) and to examine it for a bimodal distribution. "Grouped" responses should be clustering around the diagonal representing simultaneous response ($RT_1 = RT_2 + I$), while the ungrouped responses, which are expected to show long RT_2's, will be on the right-hand side of the diagonal which represents equal RT's.

There is a further characteristic which may be expected of grouped responses. Hick and Welford (1956) appear to regard the essential thing about the situation in which grouping occurs as the inability reliably to resolve the order of the two stimuli. Grouping can be expected at "all intervals less than the least at which the *order* of the two successive events is reliably distinguished; it is not sharply

defined, but is well known to be in the region of 50 to 100 millisec." One may conclude that, other things being equal, grouped responses should occur in the wrong order roughly as often as they do in the right one.

If the intermittency hypothesis is to be accepted in its present form, two nearly simultaneous stimuli must either be grouped and treated as a single situation, or one of the two must be ignored until the decision process to the other is finished. The necessity for using a single central mechanism would not in any case allow independent responses to be made in each case.

The present experiment was designed as a critical test of this prediction. The subject was presented with a situation in which the stimuli (when they both occurred) were always nearly simultaneous, so that there was every chance for him to show "grouping" of responses, while, if he did not do so, "blocking" of the second response would be pronounced and readily detectable. The two stimuli were arranged to occur within 50 or 100 millisec. of each other, and there was a high probability (0·75) of them both occurring. By using paired and unpaired stimuli at only two intervals, a great deal more intensive practice was possible than in similar experiments with a larger number of intervals, and the practice was more evenly distributed between first and second responses at any particular interval.

Method

The apparatus was similar to that used by Elithorn and Lawrence (1955). The subject sat 2·5 metres from a black bakelite screen, 61 cm. square, on which were set, 25 cm. apart, two clusters of five miniature neon bulbs. Mid-way between them was a fixation point, which could be illuminated from behind to provide a warning signal. The subject's hands rested on two keys, operating with a pressure of approximately 400 gm. Each stimulus was a 10 millisec. flash of one neon cluster, and the subject was required to respond as quickly as possible by pressing the left-hand key after the left stimulus and the right-hand key after the right stimulus.

Stimuli were presented within a period of 100 millisec.; the warning light came on 1200 millisec. before the beginning of this period and remained on for 800 millisec. The stimulus situations were presented every 4 sec., and in each situation either one or both stimuli could occur. The presentation of each stimulus and the order of presentation was decided by partial randomization with the following restrictions. Each stimulus (left and right) occurred with equal frequency at one of three times, 0, 50 or 100 millisec. after the end of the 1200 millisec. warning period, but the two stimuli were never presented simultaneously. In one situation in four either the left or right stimulus was omitted, so that only one stimulus occurred. In the paired stimulus situations there were thus six possibilities: L0/R50, R0/L50, L50/R100, R50/L100, L0/R100, R0/L100. In the first four of these situations the two stimuli are separated by an interval of 50 millisec., but the foreperiod is 50 millisec. longer in the second pair. In the last two situations the interstimulus interval is 100 millisec., and the foreperiod is the same as for the 0/50 situation. In a quarter of the situations where there was only a single stimulus there were again six possibilities: L0, R0, L50, R50, L100, R100. The order in which the various situations were presented was randomized within runs of 24 and programmed automatically by punched tape. To minimize the risk of the subject using the restricted randomization as a clue to the last few responses in each run (e.g. by counting the number of single situations), the stimuli were presented, not in blocks of 24, but in runs of 32 straddling the blocks. Daily sessions consisted of 96 situations (4 blocks) divided into 3 runs of 32 situations each, with a few minutes' rest between runs. Two subjects, both right-handed, completed 18 sessions (1,728 situations) without knowledge of results. At that time, the second subject, who had shown a most interesting bilaterally asymmetrical pattern of responses, was asked to continue the experiment. He went on to complete a total of 42 sessions (4,032 situations).

Certain abnormal responses have been excluded from the results. Premature responses can be of two kinds: those completed before the stimulus on the corresponding side and those initiated prematurely but not completed until after the stimulus appears. The latter appear as if they were short reaction times. In the present experiment all responses with reaction times less than 100 millisec. were omitted from the calculation

as being almost certainly premature. The number, which was small, is given in Table I. Similarly, the relatively few disproportionately long reaction times of more than 500 millisec. (usually associated with missing the key) have been omitted. The apparatus itself recorded premature responses made before the stimulus. These automatically stopped the presentation of further stimuli and lit one of two red bulbs located on either side of the display.

There were no designated practice periods and all the responses were recorded. Practice effects will be reported and discussed. It is worth emphasizing, therefore, that the number of responses made by our subjects (3,024 and 7,056 respectively) is high compared with that reported by most previous workers, even when practice periods are included. Moreover, this total number is distributed between a smaller number of different situations.

Results

There were three possible stimulus times for each hand, viz. immediately after the foreperiod, 50 millisec. later and 100 millisec. later (time "0," "50" and time "100" respectively). Further, at any particular time the stimulus might occur alone or with another stimulus contralaterally at either of the other two possible times. The mean reaction times to a stimulus occurring at each of the three times ("0," "50" and "100") are plotted for the left and right hands of Subject I in Figure 1 and for Subject II in Figure 2. At each time, the three possible situations (alone or with a contralateral stimulus at one of the two other times) are plotted separately. To show practice effects the means are presented for successive 6-day periods. A striking effect of this kind is seen in the period-to-period changes of mean RT_1 (Figs. 1 and 2, open symbols). In both subjects the left hand tends to start much slower in its responses than the right and to speed up later, apparently at the expense of the right hand. The latter (at least initially) becomes slower in responding and even at the end of the experiment is never appreciably faster than it was in Period 1. In contrast to this asymmetry in the first or single responses, there is much less difference between the speed of response of the two hands to a second stimulus (Figs. 1 and 2, filled-in symbols). All the second responses of Subject I tended to speed up with practice. Subject II, on the other hand, showed a sudden slowing down of his second responses, unaccompanied by any comparable change in RT_1, at the time he was told that he had to continue the experiment for a further period. This effect is similar and equally marked for both hands. The occurrence of apparently independent changes in the speed of first or second responses in the same hand is a remarkable feature of these results.

Apart from practice effects, the second subject in his longer experiment also shows some less consistent period-to-period fluctuations in mean RT, rather similar for both hands and at all the times. Both subjects show some differences in speed of response at the different times. Except for Subject II's right hand, there is a consistent tendency for RT_1 to be faster at the later times.

Performance in the double stimulus situation: RT_1 and RT_2 compared

It is clear from Figures 1 and 2 (open symbols) that there is little or no difference between the speed of response to a stimulus occurring first, whether it is subsequently followed by a second stimulus or not. This forces the conclusion that single stimuli are being treated in the same way as if they were "firsts," but does not exclude the possibility that waiting to see if a second stimulus is coming may be artificially inflating the RT in both situations. It does suggest, however, that if there is any waiting it must be for approximately the same period in all three cases. Unfortunately, the possibility of concealed waiting means that RT_1 is not a truly effective

measure of "normal" RT_2 (RT_N), and this must be borne in mind in assessing the amount of delay in RT_2. The delay may appear to be smaller than it actually is.

Since mean reaction times may not accurately reflect the pattern of individual responses, the RT's for individual pairs of responses were plotted for each 6-day period. Eight of these scattergrams are shown in Figures 3 to 5. Reaction times to the first stimulus of each pair are plotted along the Y axis, and to the second along the X axis. The responses to the corresponding single stimulus situations are shown as histograms lying along these axes. The three diagonals drawn in each scattergram represent three possible time relationships between the pairs of responses: namely, first and second responses made simultaneously, responses in which

FIGURE 1

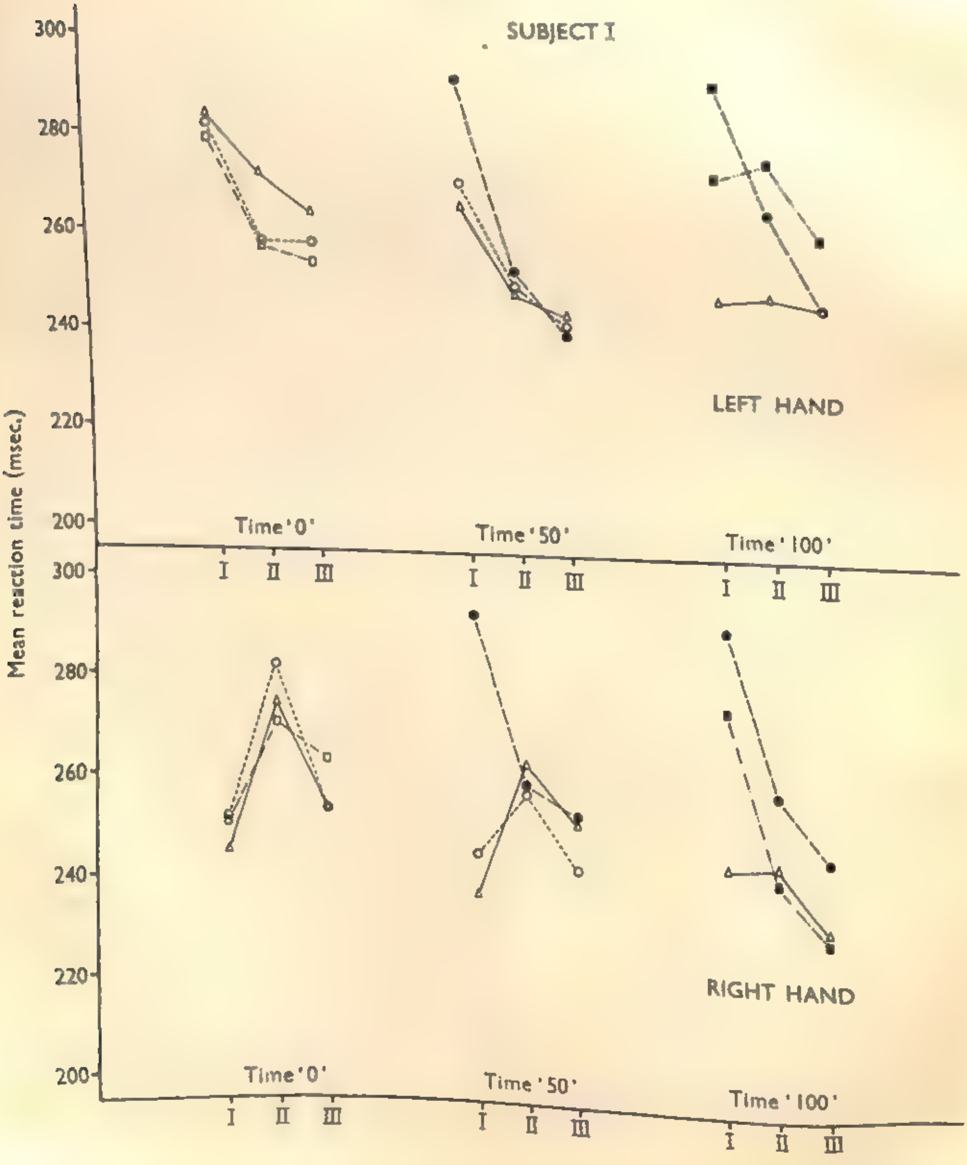
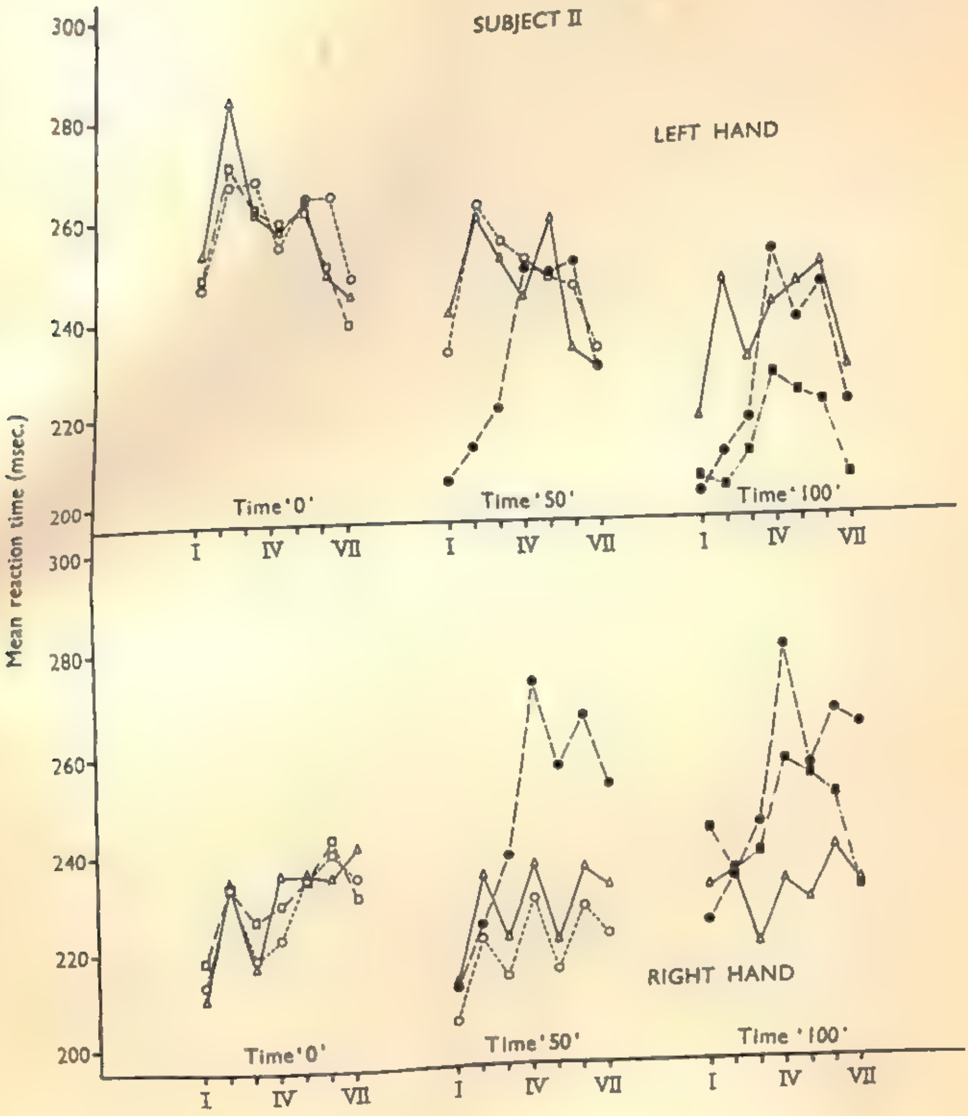


FIGURE 2



FIGURES 1 and 2

Mean reaction time of each hand for successive 6-day periods of the experiment (I, II . . ., etc.) to stimuli occurring at each of three possible times, viz. immediately after the foreperiod (time "0"), 50 millisec. later (time "50") and 100 millisec. later (time "100"). Figure 1 shows mean RT's during the three 6-day periods completed by Subject I. Figure 2 shows mean RT's during the seven 6-day periods completed by Subject II.

In Figures 1 and 2 the following conventions are adopted. Means of the responses to *single* or *first* stimuli are shown in *open* symbols; those to *second* stimuli in *filled-in* symbols. *Circles* indicate responses with an interstimulus interval of 50 millisec.; *squares*, 100 millisec.; *triangles*, a single stimulus situation.

FIGURE 3

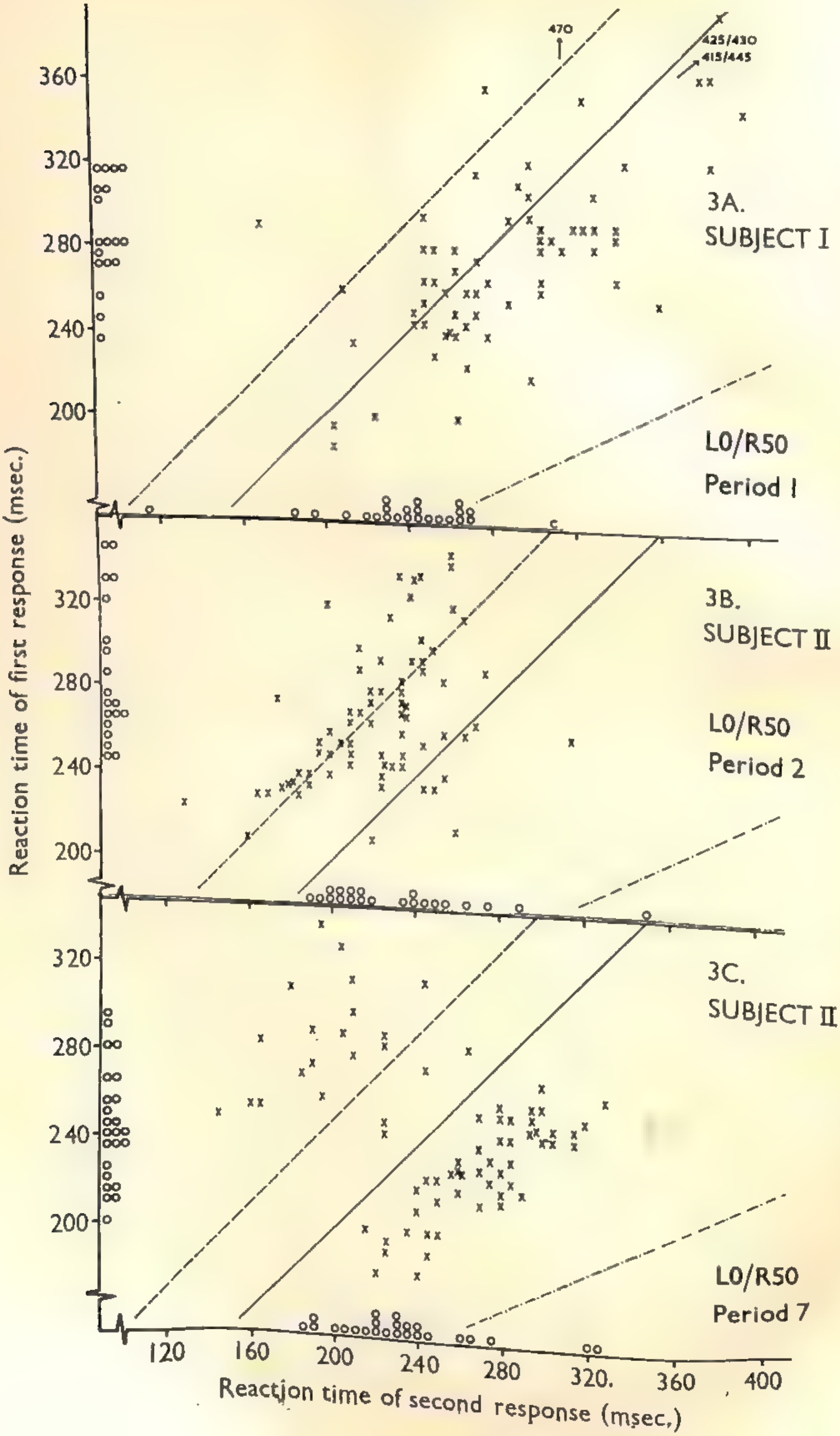

470
425/430
415/445
3A.
SUBJECT I
L0/R50
Period 1
3B.
SUBJECT II
L0/R50
Period 2
3C.
SUBJECT II
L0/R50
Period 7
Reaction time of first response (msec.)
Reaction time of second response (msec.)
120
160
200
240
280
320
360
400

FIGURE 4

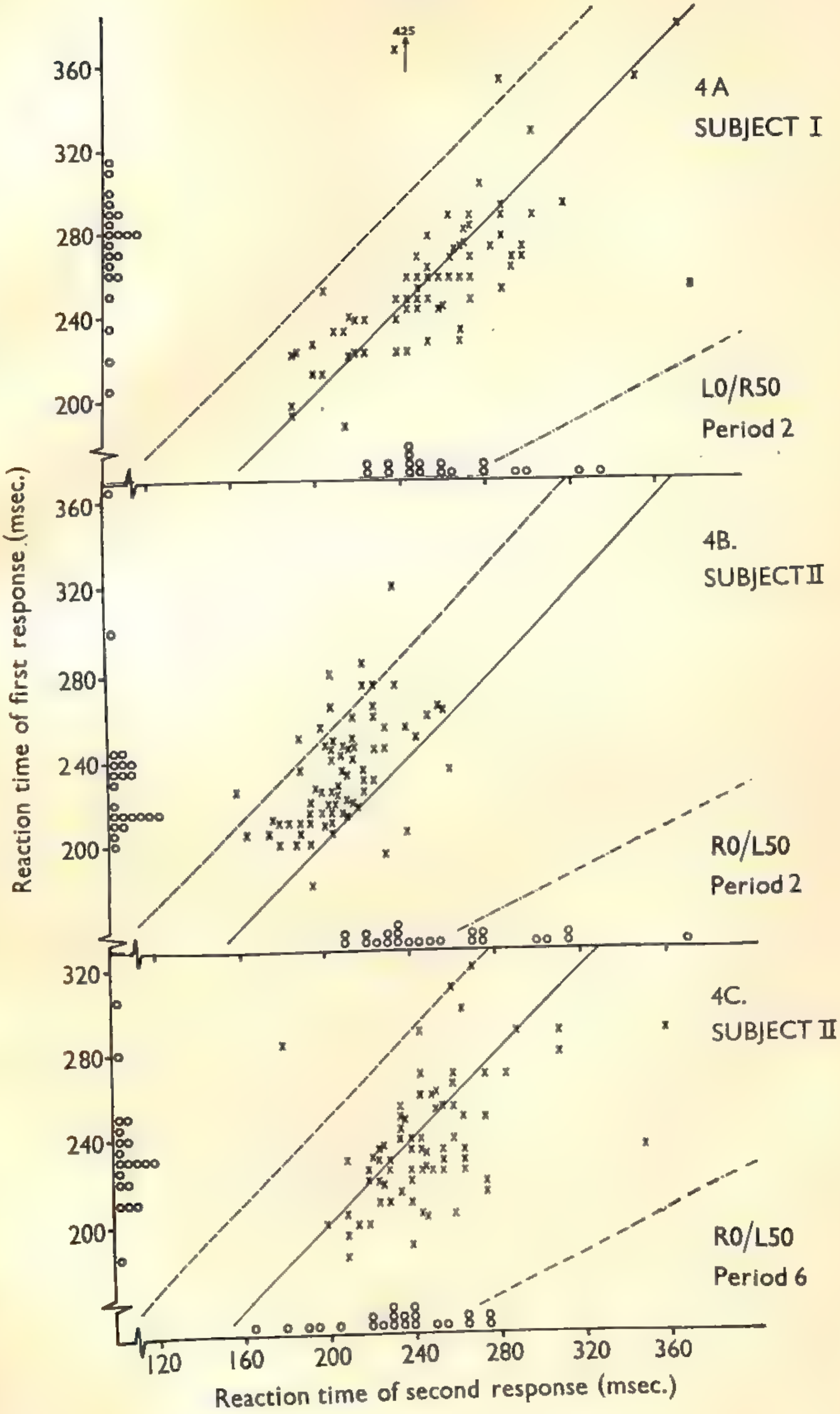
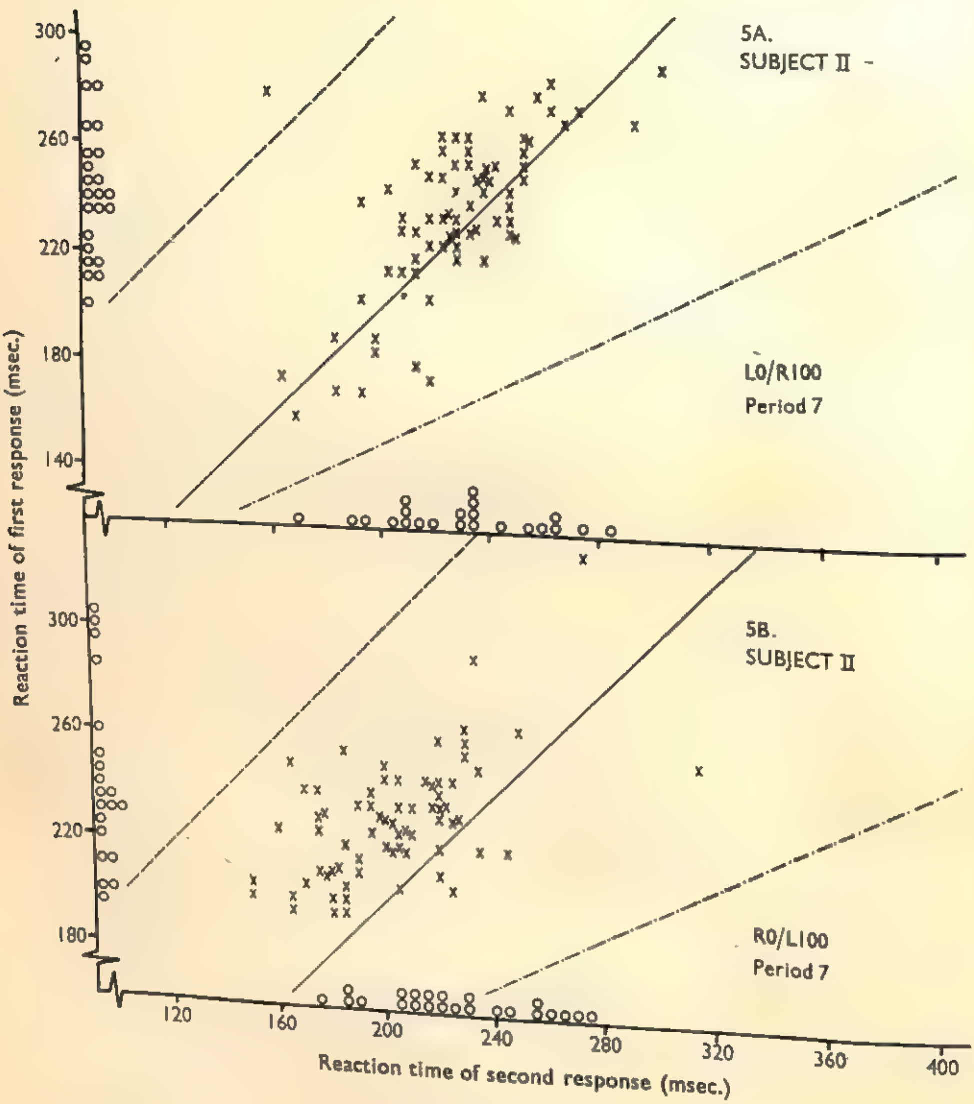

425
4 A
SUBJECT I
L0/R50
Period 2
4B.
SUBJECT II
R0/L50
Period 2
4C.
SUBJECT II
R0/L50
Period 6
360
320
280
240
200
120
160
200
240
280
320
360
Reaction time of first response (msec.)
Reaction time of second response (msec.)

FIGURE 5



FIGURES 3 to 5

Scattergrams for all individual pairs of responses in a particular 6-day period of the experiment. RT_1 is plotted along the ordinate, RT_2 along the abscissa, and the RT's to the corresponding single stimulus situation at the same time for the same hand are shown as histograms lying along the axes. The solid line diagonal indicates *equal* RT_1 and RT_2 ("independence"); the interrupted diagonal to the left of this represents an RT_1 exceeding RT_2 by the interval ("waiting and grouping"), both responses being made simultaneously; the dotted diagonal to the right of the scattergram represents an RT_2 delayed with respect to RT_1 by the amount (RT_1 − I) ("central intermittency").

$RT_1 = RT_2$ and responses in which RT_2 exceeds RT_1 by $RT_1 - I$. Grouped responses which are made nearly simultaneously will lie along the interrupted diagonal to the left, while independently organized responses should reproduce the interstimulus interval and lie close to the solid diagonal of equality. The intermittency model put forward by Davis does not permit one to predict the line about which the delayed second responses should lie, unless assumptions are made about the proportion of the reaction time which should be attributed to central processes. The third diagonal on the extreme right of each diagram therefore illustrates the relationship predicted by the intermittency hypothesis as formulated by Welford, in which RT_2 is prolonged by a delay equal to $RT_1 - I$.

The procedure adopted was to examine the means of the double responses for indications of grouping, delay or independence, and to compare these with the appropriate scattergram for confirmation. This approach yielded interesting results, and the following situations have been chosen to illustrate the varied patterns of responses which can co-exist in the same subject in different situations, or in different subjects in the same situation.

Example I

Mean RT's for Subject I in the situation Lo/R50 are respectively 282 and 293 millisec. (Period 1) and 258 and 259 millisec. (Period 2). These would suggest that RT_2 is being delayed relative to RT_1 only slightly in Period 1 and that by Period 2 the two responses have become virtually independent. Reference to the corresponding scattergrams (Figs. 3A and 4A) show that this is substantially true. In neither period are the observations far from the line of equal RT though the scatter is greater in Period 1 and several points are in excess of the single "control" responses at R50. By Period 2 many of the points are on or near the line of equality. The mean RT_2 of 259 millisec. is less than the corresponding mean for the single responses (R50 alone), which makes it unlikely that the pattern of response here is due to the introduction of a deliberate delay in RT_2. The mean RT_1 at 258 millisec. is also well below the equivalent single response (Lo alone) and is in line with RT_1's at the same interval in other situations. There seems to be considerable evidence here that responses to two stimuli at this interval have been dealt with independently.

Example 2

The mean RT's for Subject II in the same (Lo/R50) situation in Period 7 are respectively 247 and 252 millisec., which again suggests independence. The scattergram (Fig. 3C) shows, however, that the means are derived from a bimodal distribution in which about one-third of the paired responses are reversed, i.e. made with the right hand first, and the other two-thirds are subject to appreciable delay. The impression of equality given by the means is therefore misleading. This clear-cut bimodal distribution was confined to Subject II and to the latter stages of the experiment and was only found in the left/right sequence of responses. It appears in a modified form in the other 50 millisec. interval situation (L50/R100), which is not illustrated here. Earlier in the experiment (Period 2) the means are markedly different (Lo—268, R50—224) and suggest grouping. The scattergram (Fig. 3B) confirms this as almost all the responses are above the line of equal response and many are actually reversed. There is no evidence of the bimodality which develops later. Subject II obviously found it very difficult to make independent responses in this situation.

Example 3

In the right/left sequence, however, Subject II appears to be readily capable of independent organization. The mean RT's for the situation Ro/L50 in Period 6 are 239 and 248 millisec. respectively. The suggestion of independence is confirmed in Figure 4C, in which there is little grouping and a few delayed responses, but most of the observations lie along the line of equality. As in the previous example, however, the means for Period 2 (234 and 210 millisec.) imply a tendency to group responses and this is confirmed in Figure 4B, though even here many of the pairs of responses are nearer to equality than to simultaneity. With practice this subject is clearly able to solve the grouping problem in the right/left sequence and achieve independent responses, whereas the avoidance of grouping in the left/right sequence leads to the appearance of delayed RT_2's.

Example 4

The analysis so far has been confined to the 50 millisec. interval. The means for the 100 millisec. interval situation (Period 7, Subject II) show a similar final response pattern whichever stimulus comes first. RT_1 exceeds RT_2 by 27 millisec. in the Ro/L100 situation and 7 millisec. for Lo/R100. The scattergrams (Fig. 5) show that in the sequence with the greatest difference (Ro/L100) the individual responses are made in a fairly consistent pattern with RT_2 generally less than RT_1, but they are clustered much nearer the line of equality than that of simultaneity.

Performance in the double situation: the pattern of delayed responses

It is obvious from inspection of the mean reaction times that there are many occasions in which RT_2 is delayed relative to RT_1. Nowhere, however, are the delays of the magnitude predicted by Welford's hypothesis ($RT_1 - I$). Table I shows that the greatest mean delay shown by either subject in the whole experiment is 47 millisec. (Subject I, Period 1, R50/L100). The corresponding scattergram (not illustrated) shows that the distribution of responses was fairly homogeneous, with very few grouped responses so that the size of the delay is not masked by a number of artificially high RT_1's and low RT_2's. This can be checked by calculating the mean delay of the delayed responses alone. When these alone are included in the calculations, the amount of delay only increases to 61 millisec.

It is worth studying more closely the nature of these delays. The delayed responses shown in Figure 3C, previously mentioned in Example 2, form a clearly discrete group. Here, too, the delay is much smaller than would be predicted by Welford's formula, the responses being well to the left of the right-hand diagonal in the scattergram. Mean delay of these responses is actually 44 millisec. There is a close relationship between RT_1 and RT_2 ($r = 0{\cdot}79$) but not between RT_1 and the delays ($r = 0{\cdot}14$). In other words, the delay is not proportionately greater when RT_1 is long than when RT_1 is short.

Quite apart from the absolute size of the delay, there is another important discrepancy in the relative amount of delay at the different interstimulus intervals. According to all the variants of the intermittency theory this should be 50 millisec. less for an interstimulus interval of 100 millisec. than for the 50 millisec. interval. Unlike the absolute delay this *difference* should remain unaltered even if there is a constant amount of waiting in both situations altering the value of RT_1 and is not affected if some other function than total RT_1 is taken to represent the central organizing time. It should still be the same even if a substantial part of RT_1 is due to peripheral conduction time, provided that the central organizing of the first response

TABLE I

SUBJECT I

Stimulus situation	*Period*	*Mean RT for left hand*	*Mean RT for right hand*	*N**	*No. of premature responses*	*Responses outside range*	*Errors in recording*
R alone at 0	1		244·5 ± 6·4	21			
	2	–	274·1 ± 7·3	22	2	–	–
	3		253·9 ± 4·9	19	2		
L alone at 0	1	283·9 ± 5·9		18	2		
	2	272·1 ± 6·0	–	21	3	–	–
	3	263·9 ± 4·8		22	1		
R alone at 50	1		237·5 ± 7·6	24			
	2	–	263·2 ± 7·3	19	3	–	–
	3		251·6 ± 7·5	22	1		
L alone at 50	1	265·8 ± 6·8		20	4		
	2	247·6 ± 3·8	–	19	4	–	–
	3	243·9 ± 4·7		23			
R alone at 100	1		243·6 ± 6·7	25			
	2	–	245·2 ± 4·3	25	1	–	–
	3		232·9 ± 7·3	21	4		
L alone at 100	1	248·5 ± 7·6		26			
	2	249·1 ± 7·6	–	23	1	–	–
	3	247·1 ± 6·5		22	1		
R0 L50	1	292·5 ± 5·5	251·0 ± 5·1	67		2	
	2	253·6 ± 4·3	282·1 ± 7·5	70	–	2	
	3	240·1 ± 3·3	253·0 ± 4·3	70		1	1
L0 R50	1	282·0 ± 6·4	293·0 ± 6·5	67		2	1
	2	257·6 ± 5·0	259·4 ± 4·7	71	–	2	
	3	257·9 ± 3·6	252·9 ± 4·0	71			1
R50 L100	1	291·6 ± 5·8	244·9 ± 3·9	63		1	
	2	265·8 ± 4·9	256·8 ± 4·0	71	–	1	
	3	247·5 ± 5·6	242·3 ± 3·6	66		1	2
L50 R100	1	271·3 ± 7·1	291·1 ± 7·8	66		2	1
	2	250·1 ± 3·3	259·4 ± 4·2	69	–		1
	3	242·1 ± 3·5	247·0 ± 4·5	68		1	
R0 L100	1	273·4 ± 6·1	249·7 ± 2·7	70			
	2	276·5 ± 5·9	270·4 ± 4·2	68	–	3	1
	3	261·0 ± 8·2	263·6 ± 4·0	66		1	3
L0 R100	1	279·4 ± 4·3	274·7 ± 5·4	71	1		
	2	257·1 ± 3·2	241·8 ± 5·8	68		2	1
	3	253·9 ± 2·8	231·1 ± 4·9	66		1	

* Number of responses used in all calculations, excluding those in the following three columns.

TABLE I (*contd.*)

SUBJECT II

Stimulus situation	*Period*	*Mean RT for left hand*	*Mean RT for right hand*	*N**	*No. of premature responses*	*Responses outside range*	*Errors in recording*
R alone at 0	1		210·0 ± 7·1	16	7		1
	2		233·9 ± 7·7	22	2		
	3		216·3 ± 4·1	20	2		2
	4	–	235·7 ± 6·2	22	2	–	
	5		235·5 ± 6·4	21	3		
	6		233·8 ± 5·7	20	4		
	7		240·0 ± 7·5	20	4		
L alone at 0	1	254·2 ± 9·5		20	3		1
	2	285·0 ± 7·6		19	5		
	3	262·5 ± 6·1		22	2		
	4	258·4 ± 4·8	–	22	2	–	
	5	264·0 ± 5·3		20	4		
	6	248·1 ± 7·1		24			
	7	243·5 ± 5·3		24			
R alone at 50	1		211·8 ± 8·8	22	2		
	2		233·8 ± 8·2	21	3		
	3		221·5 ± 5·5	20	4		
	4	–	236·1 ± 6·2	22	2	–	–
	5		219·8 ± 6·2	23	1		
	6		226·3 ± 9·5	23	1		
	7		231·3 ± 6·9	24			
L alone at 50	1	238·1 ± 8·7		21	3		
	2	258·0 ± 8·3		23	1		
	3	249·3 ± 6·0		21	1		2
	4	245·2 ± 4·9	–	24		–	
	5	256·7 ± 6·2		24			
	6	230·2 ± 6·1		22	2		
	7	235·5 ± 7·0		21	3		
R alone at 100	1		229·7 ± 7·7	19	5		
	2		232·9 ± 7·5	22	2		
	3		218·3 ± 5·3	23	1		
	4	–	230·9 ± 8·3	23	1	–	–
	5		228·3 ± 5·9	24			
	6		238·3 ± 7·1	23	1		
	7		230·7 ± 6·1	22	2		
L alone at 100	1	214·7 ± 8·4		20	4		
	2	242·9 ± 6·7		22	2		
	3	231·4 ± 5·6		21	3		
	4	237·9 ± 7·5	–	24		–	–
	5	238·0 ± 4·9		23	1		
	6	246·0 ± 9·6		24			
	7	225·4 ± 5·9		23	1		

* Number of responses used in all calculations, excluding those in the following three columns.

TABLE I (*contd.*)

SUBJECT II

Stimulus situation	*Period*	*Mean RT for left hand*	*Mean RT for right hand*	*N**	*No. of premature responses*	*Responses outside range*	*Errors in recording*
Ro L50	1	203·2 ± 3·8	213·1 ± 3·3	70	1		1
	2	210·3 ± 2·4	233·5 ± 2·7	72			
	3	217·8 ± 3·7	218·5 ± 3·6	71			1
	4	247·0 ± 3·3	222·4 ± 2·8	72		–	
	5	246·0 ± 4·4	233·7 ± 3·7	72			
	6	247·9 ± 3·5	239·2 ± 3·5	72			
	7	226·2 ± 3·2	233·9 ± 3·7	72			
Lo R50	1	246·9 ± 4·3	211·3 ± 4·9	72			
	2	267·7 ± 3·6	224·1 ± 3·7	72			
	3	268·8 ± 4·2	237·8 ± 5·7	69		1	2
	4	255·1 ± 4·1	273·3 ± 5·6	71	–		1
	5	265·3 ± 4·3	255·6 ± 5·8	72			
	6	265·1 ± 5·9	266·3 ± 4·6	72			
	7	246·9 ± 4·1	252·0 ± 5·3	71			1
R50 L100	1	200·3 ± 3·5	204·5 ± 3·4	71	1		
	2	208·2 ± 2·7	220·7 ± 3·4	72			
	3	214·6 ± 3·3	213·1 ± 3·2	71			1
	4	248·9 ± 3·7	228·8 ± 4·0	70	1	–	1
	5	235·2 ± 3·8	213·7 ± 4·3	72			
	6	242·4 ± 4·8	226·8 ± 4·1	72			
	7	218·2 ± 3·9	221·5 ± 2·6	72			
L50 R100	1	228·7 ± 3·8	223·2 ± 5·1	71			1
	2	261·1 ± 5·0	232·2 ± 3·9	72			
	3	253·5 ± 4·3	243·5 ± 5·7	72			
	4	248·6 ± 3·8	278·7 ± 5·9	71		–	1
	5	245·1 ± 3·9	255·3 ± 5·0	71			1
	6	242·8 ± 4·8	265·6 ± 4·8	72			
	7	229·8 ± 2·8	262·8 ± 5·0	71	1		
Ro L100	1	203·4 ± 5·2	218·2 ± 4·7	71			1
	2	201·4 ± 3·9	233·2 ± 4·4	70	1		1
	3	208·4 ± 3·8	226·4 ± 4·1	67		1	4
	4	224·4 ± 3·7	229·4 ± 3·6	72			
	5	219·6 ± 4·1	234·0 ± 4·9	72			
	6	218·3 ± 3·9	242·2 ± 3·8	72			
	7	203·0 ± 3·5	229·9 ± 3·1	70	2		
Lo R100	1	249·2 ± 3·8	241·7 ± 3·1	68	2	1	1
	2	272·0 ± 3·7	232·7 ± 3·9	69	1	1	1
	3	263·4 ± 3·7	237·2 ± 3·6	71	1		
	4	259·7 ± 3·4	256·3 ± 3·4	72			
	5	262·3 ± 4·2	252·6 ± 3·8	71			1
	6	250·5 ± 3·8	249·6 ± 4·5	71			1
	7	237·2 ± 3·4	229·9 ± 2·8	72			

* Number of responses used in all calculations, excluding those in the following three columns.

time is not less than 100 millisec. A further factor must be considered. As Elithorn and Lawrence (1955) pointed out, the expected RT_2 cannot strictly be inferred from the observed RT_1, since the arrival of the first stimulus changes the subject's set in relation to the second. The arrival of S_1 converts the originally disjunctive situation into a simple response. Leonard (1958) has confirmed that advance information of this type can be used more effectively to speed up a reaction time to a second stimulus as the interval between stimuli increases. This means that the expected RT_2 will be reduced as the interval increases. In the present case this would increase the difference between the delays shown at the 50 and 100 millisec. intervals, i.e. $(RT_2 - RT_1)_{50} - (RT_2 - RT_1)_{100}$. A comparison of these delays shows that the reduction at the longer intervals is often much less than the 50 + millisec. predicted by the combination of Welford's theory and Leonard's observation. The observed differences range from a minimum of 0 millisec. (Subject II, Period 6: L0/R50 and L0/R100) to a maximum of only 40 millisec (Subject II, Period 7: L50/R100 and L0/R100).

Discussion

The data here presented show beyond question that even with two stimuli separated by only 50 or 100 millisec. it is possible to make second responses which are as fast as those made to a single stimulus, but are not performed almost simultaneously with the first response, as would be expected were they being grouped. Before accepting the view that these undelayed second responses are truly independent (in the sense of being made separately, in response to the second stimulus, after the central processes concerned in the first response are already under way) two possibilities need to be considered, viz. (1) that they are premature responses, initiated before the arrival of the stimulus, but masked by its subsequent occurrence, and masquerading as genuine responses shorter than they really are; and (2) that they represent a special kind of concealed grouping, in which the responses are made, not simultaneously, but one after the other in the right order and at the right interval, as the outcome of a single "decision process."

The design of the experiment might have been expected to lead to an appreciable number of premature responses, since stimulus probability was high over a very short range of time and guesses as to which side the stimulus was coming have never less than a 50 per cent. chance of being right. However, an index of the percentage of premature responses occurring is given by the number recorded on the 25 per cent. of occasions when there was no stimulus on one of the two sides. The number and distribution of these is recorded in Table I: Subject I had 29 in 421 situations (6·9 per cent.) and Subject II 86 out of 1008 (8·5 per cent.). Premature responses may be made on a particular side either before S_1 has arrived or after a contralateral S_1 in anticipation of an S_2. Since both these types of premature response contribute to the percentage figures for the single stimulus situation calculated above, it seems reasonable to assert that less than 9 per cent. of the second responses on each side can be premature. Such a percentage clearly cannot account for the very large proportion of undelayed second responses, performed with a reaction time almost equal to that of the first response.

The second possibility is more difficult to exclude, but there are reasons for regarding it as equally unlikely. The suggestion is that the subject might be grouping the responses in a rather sophisticated way, and deciding to do them, not simultaneously, but in the correct sequence and at an interval corresponding to that between the two stimuli. As only a single decision process is involved, the intermittency hypothesis is still satisfied, although grouping can no longer be regarded

as something which occurs when the order of the two stimuli cannot be reliably distinguished. In the pattern of undelayed second responses the order and the interstimulus interval are not only distinguished but accurately reproduced (e.g. Figs. 4A and 5A).

In the model proposed by the intermittency hypothesis, the single decision process to respond to both stimuli cannot begin until the second stimulus has arrived. It is supposed that the grouping decision may be either to respond to both stimuli at the same time, or to respond to them consecutively at the interstimulus interval. In the first case RT_1 must have the interstimulus interval added to it; and in the second, both RT_1 and RT_2 will be increased by this amount.

As RT_1 must exceed the true reaction time by the interstimulus interval, it can only be equal to its control RT_1 (i.e. any *un*grouped first response) if there is *always* the same amount of waiting whether the responses are grouped or not. In the present experiment there is no appreciable difference in the mean RT_1's whether a second stimulus follows in 50 or 100 millisec. or not at all. Clearly, then, the amount of waiting (if any) must be the same in all cases. The subject could produce the supposedly grouped pattern of equally fast responses with stimuli 100 millisec. apart (Fig. 5), but there is no evidence that his RT_1 was quicker when both stimuli had already arrived by time "50." Assuming that grouping is occurring, it is difficult to explain why it is not 50 millisec. shorter in the latter case unless he is waiting a further 50 millisec. after *both* stimuli have arrived.

Further, if the subject always waits 100 millisec. before initiating his response, his true reaction time for both grouped responses (subtracting 100 millisec. waiting) would be of the order of 100 to 150 millisec. and at the 100 millisec. interstimulus interval for Subject II about 110 millisec. (Fig. 5). Even a simple visual reaction time is of the order of 150 millisec. and the present situation is a disjunctive one Moreover, if the grouping strategy is adopted, the situation becomes at least a three-choice one, and theory predicts that the reaction time should be considerably longer (Hick, 1952; Hyman, 1953; Welford and Hick, 1956). Some support for these doubts is also provided by an experiment of Hick (1948) in which the subject was instructed to try and group his responses. In the responses regarded by Hick as grouped, RT_2 was almost always longer than it was for the ungrouped second responses, even though refractory delays were occurring in the latter. The mean "complex" reaction time was about 430 millisec. as against 389 millisec. for the ungrouped.

In the present results, on the other hand, RT_2 is considerably *faster* than its control when the supposedly grouped patterns of responses are occurring. This applies both to the mean RT_2 (e.g. Subject II's left hand, Fig. 2) and the individual second responses (Figs. 4B and 3C). RT_1 in these responses, far from having the interstimulus interval added to it by waiting, is no longer than its own control or of the control RT_2. Instead of RT_1 being longer by the interstimulus interval and RT_2 equal to its control, RT_1 is unchanged and RT_2 has been speeded up. This may, of course, represent some degree of waiting in all the responses to S_1, but it is evident that any such waiting is in no sense peculiar to the periods when the subject is producing nearly simultaneous first and second responses, and it can hardly be of the order of 100 millisec. in all cases.

These results therefore appear to be incompatible with the suggestion that all apparent exceptions to the lengthening of the early second reaction time predicted by the intermittency hypothesis are due to grouping of responses. A large number of those second responses were not delayed, but were performed as fast as the first responses, and their timing makes it almost impossible that they can be the outcome

of a single decision process as envisaged by the single information channel model. Where the two responses were made nearly simultaneously, the second is faster than its control, and where they were made in the right order at the right interval, the second was roughly equally fast as its control; in neither case was RT_1 significantly different from its control, and its low value precludes the possibility that much waiting was occurring.

There remain the delayed second responses to be considered, and these, too, do not appear to fit easily into the framework of the intermittency theory. There are two reasons for this. Firstly, the delays were never of anything like the magnitude predicted by the $RT_1 - I$ relationship (even if due allowance is made for the conduction time in the motor pathways), and, secondly, they were largely learnt out by one of the two subjects in the course of the experiment. It seems that there must have been something in the present experimental situation peculiarly favourable to this, because it has been the experience of earlier workers that these delays are resistant to training (Hick, 1948; Davis, 1956, 1957; Marill, 1957). But Telford (1931) and Elithorn and Lawrence (1955) have previously noted that they were reduced by practice and this has recently received strong support from a study by Mowbray and Rhoades (1959), who have shown that the initially greater RT in a four-choice situation as compared with a two-choice one can be eliminated with sufficient practice. The present experiment differs from many of these earlier ones in the much greater probability of S_2 occurring 50 or 100 millisec. after S_1 and the much more intensive practice which the subjects had in responding at these times. This results from the choice of only two interstimulus intervals; where a larger range is used, each has to be presented and there will inevitably be a proportionately lower expectancy of S_2 occurring at any particular interval after S_1. It seems a reasonable surmise that the high expectancy for S_2 may account for both the shortness of the delays observed and their disappearance with training.

Welford (1952) considered the possibility that the delays in early RT_2's might be due to the unexpectedness of S_2 soon after S_1; his main reason for rejecting it was the belief that they were not reduced by practice. Hick (1948) also came to the conclusion that there were irreducible delays associated with the second of two closely-following responses. The present results, however, show that they are not inevitable, and also that they do not necessarily have an all-or-none character.

There are thus many important features of the pattern of responses in these experiments which make any attempt to fit them into the intermittency hypothesis somewhat of a Procrustean labour. This is by no means to deny that there is some form of interaction between the two central processes concerned in the two responses; it is only to deny that such interaction is limited to their being either combined in a single process (grouping) or whichever is first blocking the second until it is finished (intermittency). The present findings point to a central mechanism of considerably more plasticity. The second response, even when called upon within a few milliseconds of the first, may be apparently speeded up, slowed down or made at the same speed. The actual pattern was not necessarily the same on both sides and seemed to depend in our subjects on the *order* in which the two sides were called upon to respond.

It may be concluded that, whatever the explanation of the delayed responses upon which the idea of the "psychological refractory period" is based, the latter is not a necessary or invariable phenomenon. There is, in fact, strong evidence against any inevitable intermittency in the central mechanisms concerned with the responses to two nearly simultaneous stimuli.

Our thanks are due to Miss J. Mott and Miss R. Mingay for their help in carrying out the experiments and preparing the results for publication, and to Dr. E. A. Carmichael for his support and encouragement.

Note: While this paper has been in press, a further paper by Welford has appeared (this *Journal*, **11**, 193–210), which contains a detailed discussion of grouping of stimuli and the single channel hypothesis.

References

CRAIK, K. J. W. (1947). Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.*, **38**, 56–61.

CRAIK, K. J. W. (1948). Theory of the human operator in control systems. II. Man as an element in a control system. *Brit. J. Psychol.*, **38**, 142–8.

DAVIS, R. (1956). The limits of the "psychological refractory period." *Quart. J. exp. Psychol.*, **8**, 24–38.

DAVIS, R. (1957). The human operator as a single channel information system. *Quart. J. exp. Psychol.*, **9**, 119–29.

ELITHORN, A., and LAWRENCE, C. (1955). Central inhibition—some refractory observations. *Quart. J. exp. Psychol.*, **7**, 116–27.

HICK, W. E. (1948). The discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, **1**, 36–51.

HICK, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11–26.

HICK, W. E., and WELFORD, A. T. (1956). Comments on Elithorn and Lawrence (1955). *Quart. J. exp. Psychol.*, **8**, 39–41.

HYMAN, R. (1953). Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, **45**, 188–96.

LEONARD, J. A. (1958). Partial advance information in a choice reaction task. *Brit. J. Psychol.*, **49**, 89–96.

MARILL, T. (1957). The psychological refractory phase. *Brit. J. Psychol.*, **48**, 93–7.

MOWBRAY, G. H., and RHOADES, M. V. (1959). On the reduction of choice reaction times with practice. *Quart. J. exp. Psychol.*, **11**, 16–23.

TELFORD, C. W. (1931). The refractory phase of voluntary and associative responses. *J. exp. Psychol.*, **14**, 1–36.

VINCE, MARGARET A. (1948). The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, **38**, 149–57.

WELFORD, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance—a review and a theory. *Brit. J. Psychol.*, **43**, 2–19.

STUDIES IN OLFACTORY ACUITY. I.

Measurement of Olfactory Thresholds in the Rat

BY

J. T. EAYRS* AND D. G. MOULTON†

From the Department of Anatomy, University of Birmingham

The capacity of the rat to respond to olfactory cues has been studied in three experimental situations: (i) an elevated Y-maze, (ii) a rectangular choice apparatus in which extraneous environmental cues were present, and (iii) a circular choice apparatus from which such cues were either excluded or randomized.

No rat showed any sign of learning on the Y-maze even after extensive training and the application of electric shock as deterrent for error.

Some rats succeeded in performing an olfactory discrimination in the rectangular choice apparatus but learning was retarded, and in some cases inhibited, by the counter-influence of non-olfactory factors.

All rats readily learned an identical discrimination when shielded from such cues in the circular choice apparatus which provides a satisfactory medium both for training and for the study of olfactory acuity.

Differences in the behaviour of rats on these three types of apparatus suggest first, that the formation of learned habits based on olfaction depends on the close contiguity of stimulus and reward, and second that the sense of smell possesses weak orientating properties by comparison with those exerted by other modalities. The findings thus help to explain differences in the ease or difficulty with which olfactory discriminations have been established by earlier workers.

Introduction

The electrophysiological studies of Adrian (1951) together with the work of Guillot (1948) on partial anosmia have suggested that there may be a number of functionally distinct types of olfactory receptor and this view has recently received support on histological grounds (Clark, 1956). But although these findings provide a basis for understanding the way in which the olfactory organ can discriminate between different odours the mechanisms by which odorous substances excite the various olfactory receptors is still far from clear.

A possible approach to this problem is through an analysis of the relationship between the physico-chemical properties of odorants and their relative detectability. Most earlier work on these lines has been confined to a study of olfactory thresholds in man and insects and little attempt has been made to exploit the possibilities of macrosmatic laboratory animals as test subjects. A probable reason for this, at least as far as the rat is concerned, is the obvious difficulty experienced by previous investigators (e.g. Liggett, 1928; Swann, 1933; Brown and Ghiselli, 1938) in developing satisfactory techniques for establishing habits based on olfactory discrimination, for only when the odour of food itself has provided a major component of the positive stimulus has rapid learning been obtained (French 1940; Stone, 1941; Lashley and Sperry, 1943).

It is thus clear that methods which have proved satisfactory for teaching discriminations based on non-olfactory modalities of sensation are unsatisfactory for quantitative studies on olfaction either by reason of the difficulty with which they

* Henry Head Research Fellow of the Royal Society.
† Present address: Division of Physiology, Florida State University, Tallahassee.

can be successfully applied or of the restricted nature of the stimuli to which the rat will readily respond. Any alternative must not only overcome these disadvantages but must also ensure that all possibility of response to extraneous cues is eliminated. The methods, for example, of Le Magnen and Rapaport (1951), who trained rats to discriminate between drinking bottles on the basis of odorous substances smeared on to their spouts, and of Gruch (1957) whose rats chose one of three boxes containing food on the basis of odorous or non-odorous air admitted from an associated tube, may be criticized on the ground that visual cues and the smell of the reward were not adequately controlled. The sensitivity of the rat to the odour of food is in fact so keen (Hand, 1952; Le Magnen, 1952) as to effectively rule out its use as a direct reward in experiments designed to estimate olfactory thresholds; for even if controlled (as, for instance, by their presence in relation to both positive and negative stimuli) odorous food molecules may mask the presence of the test substances either by competing for olfactory receptors or by chemical reaction.

The purpose of the present paper is to describe a method for the study of olfactory acuity in the rat* and, since certain preliminary observations would appear to throw some light on difficulties experienced by earlier workers, to outline the several stages leading to its development. The fuller use of this technique for demonstrating the olfactory thresholds of a series of chemically related compounds is described in the following paper.

Material and Methods

Apparatus

Three different sets of apparatus, designated the (i) Y-maze, (ii) rectangular, and (iii) circular choice boxes were used during the course of the experiments. The initial design and incidental modification of this apparatus were based on experimental observations made during its use and are therefore more conveniently described together with the Results.

Animals

The preliminary experiments reported here were carried out using albino rats of the Birmingham strain. In view of claims that olfactory acuity in man is influenced by the ovarian cycle (Elsberg, Brewer and Levy, 1935; Le Magnen, 1950; Schneider and Wolff, 1955) only male rats were used.

Odorants

The odorants used during these preliminary trials were oil of cloves (eugenol) and hexanol. These were initially dissolved at a concentration of 10 per cent. in propylene glycol, subsequent aqueous dilutions being made by factors of 10. Equivalent dilutions of propylene glycol alone acted as controls. A further series of tests was carried out using aqueous ethanol. No special precautions were taken to eliminate possible contamination.

Mode of estimating thresholds

(*a*) *Conduct of tets.* For reasons described later only the circular choice apparatus was ultimately used to measure thresholds. Rats were motivated by restricting their water intake to about 15 ml. daily and were introduced into the apparatus manually at each trial. Where, as in later trials, the odour to be detected was carried in nylon capsules, these were discarded after use so that no rat was presented with the same capsule more than once.

After preliminary training the approximate threshold for each test substance was determined by conducting a preliminary series of trials based on graded concentrations

* Since the work reported in this paper was completed Pfaffmann, Goff and Bare (1958) have published a technique based on operant conditioning which would appear to meet most of the methodological criticisms outlined and to serve the same end as the method here described.

of odorant ranging from 10^{-1} to 10^{-14} and analysing the data by the method of probits. The range of concentrations ultimately used for testing was centred about the value so obtained for a 50 per cent. success score. Rotation of rats in relation to the sequence of trials resulted in the intervention of an interval of at least 30 sec., and usually longer, between trials for any given rat and this was thought sufficient to minimize the possible effects of adaptation to the odour. Further, in order to minimize the influence of day-to-day fluctuations in performance the concentrations chosen were presented in accordance with a balanced incomplete block design in which all concentrations were replicated the same number of times. The designs used (see Fisher and Yates, 1948) were the 9×12 (i.e. 9 dilutions at the rate of 4 per day over 12 days with 3 replications), 7×7 (4 dilutions per day with 3 replications) or 16×16 (6 dilutions per day with 6 replications).

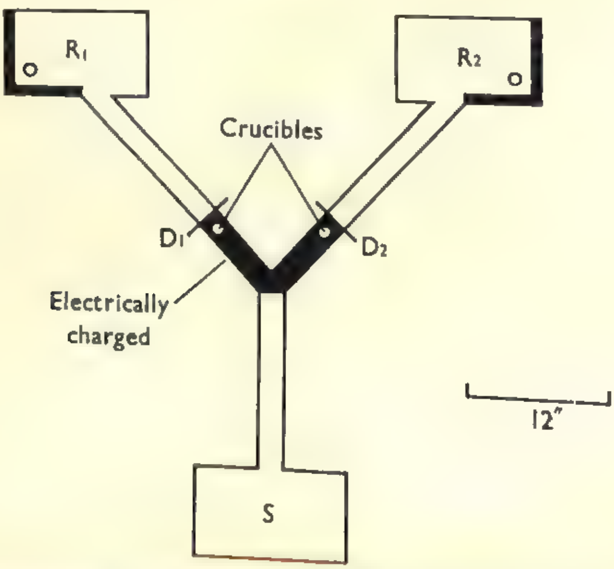
(*b*) *Scoring and treatment of data.* The number of correct and incorrect choices was recorded, and on completion of each block of trails the rectified mean percentage of correct responses at each concentration was computed in accordance with the usual procedure for balanced incomplete blocks (Yates, 1936) and the values so obtained subjected to probit analysis (Finney, 1952). A separate probit regression was computed for each set of data and used to obtain the olfactory threshold in terms of the concentration detectable at the 50 per cent. level. The thresholds thus obtained were expressed as the dilution of the odorant in solution, and since the object of these preliminary experiments was merely to establish the usefulness of a technique, the data were subjected to no further treatment. In later experiments, however, in which a comparison of the relative detectability of different odorants has been attempted, more refined procedures have been adopted (Moulton and Eayrs, 1960).

Results

(*a*) Y-*maze*

The apparatus used (Fig. 1) was similar in general construction to the 3-table apparatus of Maier (1932). A crucible, covered by a copper gauge was sunk into each of the two elevated runways leading to the reward tables (R1, R2), one containing a strong solution of oil of cloves, the other the solvent alone. Water contained in a dish hidden from sight was available on one of the reward tables; an empty dish was similarly located on the other. The right/left positioning of the crucible containing the oil of cloves and positive reward table were randomly interchanged.

FIGURE 1



Y-maze—schematic diagram showing spatial relationship of components.

Six thirst-motivated rats, which had previously been habituated to the apparatus by free exploration, were each given 30 trials a day in two sessions of 15 trials each. For three of these rats oil of cloves marked the runway leading to the positive reward table; for the other three it acted as negative stimulus. Each rat was placed on the start-table (S) and if choosing the correct pathway was allowed to drink for about 5 sec. before being returned to the start-table for the next trial. In instances where rats retraced their path an error was scored if the forepaws had previously crossed the position occupied by the negative crucible, but in practice use of this criterion was rarely necessary since rats, once started along a given pathway, almost invariably continued to reach the reward table.

During the first 300 trials the rats were allowed to retrace after error and to find their way to the correct table; for a further 300 trials the animals were immediately removed from the negative table and replaced on the start table for the next trial. At the end of these two series (600 trials per rat) no rat showed any evidence of learning, the correct and incorrect runways being chosen in approximately equal numbers. It seemed possible that remoteness of stimulus and reward might account for the rat's failure to learn and reversible one-way doors (D1, D2) were accordingly fitted on the immediate far side of the odour-containing crucibles. The door on the positive runway was arranged to swing outwards and so permit access to water on the reward table; that on the negative runway arrested further progress in a position where the rat could scarcely fail to detect the presence or absence of the olfactory stimulus. An additional 300 trials was run, any attempt to push against a door on the negative runway being scored as an error. There was no evidence of learning at the end of this period; nor did the administration of an electric shock as punishment for pushing the negative door prove effective during the course of a further 300 trials.

(*b*) *Rectangular choice box*

This apparatus consisted of an elevated runway leading to a choice box at the far end of which were hung two water bottles. The drinking spouts of these bottles were modified by the attachment of a short length of blind-ended tube just above the nozzle into which pledgets of cotton wool soaked in odorant or control solution could be placed. The spouts were insulated from the wire cage by rubber tubing, that containing the negative stimulus being connected to a source of potential about 45 V. above that of the floor of the cage. In drinking from this tube a rat would thus receive a shock derived from an estimated current of about 1 mA.

Six thirst-motivated rats were taught initially to run the length of the runway to reach a single central drinking spout. Later the single spout was replaced by the negative (odour charged) and positive (odour free) spouts whose relative positions to right and left of the choice box were varied in accordance with a Gellerman (1933) sequence. All six rats were run in randomized order for any one setting of positive and negative drinking tubes after which the tubes were reset for the next series of trials.

The behaviour of the rats in this experimental situation precluded the running of tests based on a fixed number of trials daily as had been used in the previous experiment. On receiving a shock as a result of choosing of the negative tube all rats refused to drink and during later tests, when motivation was raised, the rats' behaviour was characterised by the appearance of innately organized responses of the type seen in displacement reactions such as scratching, face-washing, leaping and hopping and, in several instances, attempted copulation by one male upon another placed in the choice box. These activities were interspersed with constant approaches and withdrawals to one or other of the drinking tubes. Where the tube finally sampled

proved to be odour free (positive) the rats rapidly learned to drink from it in subsequent trials, all displacement activity disappearing. When, however, the positions of positive and negative tubes were interchanged all rats attempted to drink from the tube occuping the previously positive position rather than responding to the associated odour and, on receiving a shock, exhibited the identical range of behaviour seen during earlier training. Two of the six rats used, however, eventually learned to respond to the odour rather than to the position of the drinking bottles.

These observations suggested that the rat might satisfactorily be taught to perform olfactory discriminations by a method basically similar to that described provided that the time-consuming obstacles arising from (*a*) the initial fear of the apparatus and (*b*) the conflict between two incompatible orientating cues could be overcome. To this end, the apparatus described in the following section was devised in which the first of these difficulties was met by arranging for training to be carried out mechanically without requiring the presence of the operator and the second by eliminating as far as possible all guides to orientation other than could be derived from olfactory stimuli.

(*c*) *Circular choice apparatus*

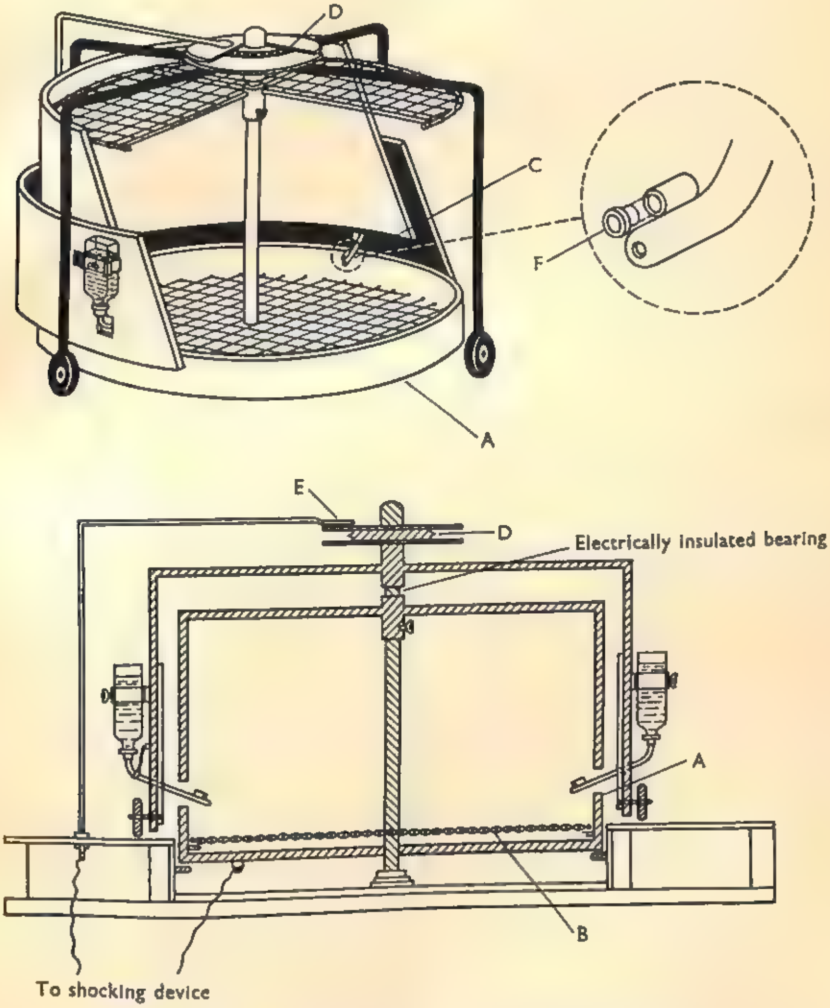
The apparatus finally developed is illustrated in Figure 2 and consists essentially of two concentric sheet metal cylinders, the outer of which is free to rotate about the inner and is electrically insulated from it. The lower part of the inner cylinder consists of a circular tray (A) containing sawdust which is provided with a copper mesh grid (B) on which the rat stands. The upper part of this cylinder, suspended from a central pillar by a spider, is separated from the tray by a 1 in. gap through which the spouts of drinking bottles (C) attached to the outer cylinder protrude into the inner compartment. The roof of this compartment is closed with steel mesh, a door being provided to permit ready access. The outer cylinder is pivoted by means of an insulated bearing, about the central pillar and rides on rubber-tyred wheels. The spider supporting this component is surmounted by a sprocket and brass plate (D). The former enables the outer cylinder to be driven slowly around the inner by means of a chain drive from a similar sprocket fitted to the spindle of a kymograph. A copper foil brush (E) rides on the brass plate and this and the inner cylinder (and hence the copper mesh floor) are connected to the opposite poles of an electrical shocking device. The drinking tubes are similar to those described for the previous apparatus except that the impregnated cotton wool pledgets are carried in a detachable nylon capsule (F). The negative tube is connected to the outer cylinder by means of a wire and clip invisible to a rat inside the apparatus.

Preliminary training is carried out by fitting each drinking tube with a capsule containing the test odorant or control solution and connecting up the electric circuit appropriately. One or two thirsty rats are placed in the apparatus which is set to rotate at about 1 revolution in 12 hr. Only cues arising from the odour of the tubes are available to the rat, for the positions of the tubes relative to the living compartment and to the room in which the apparatus is housed are being constantly changed and the possibility of visual cues is reduced by keeping the apparatus in darkness. Later training is carried out manually, the rat being introduced into the apparatus for each trial and the positions of the drinking spouts set at random by rotating the outer cylinder between individual trials. The rat's success or otherwise is indicated by means of a trigger unit linked to the shocking system which illuminates a neon bulb each time the rat makes an error. Successes are scored for a constant and readily detectable concentration of the odorant until a plateau

is reached when the animal is regarded as ready for determination of the olfactory threshold for this substance.

The ease with which rats could be trained in this apparatus fully justified its design. Rats left in the apparatus for two nights were usually fully habituated to the experimental situation and, after 24 hr. without water, would, without further

FIGURE 2

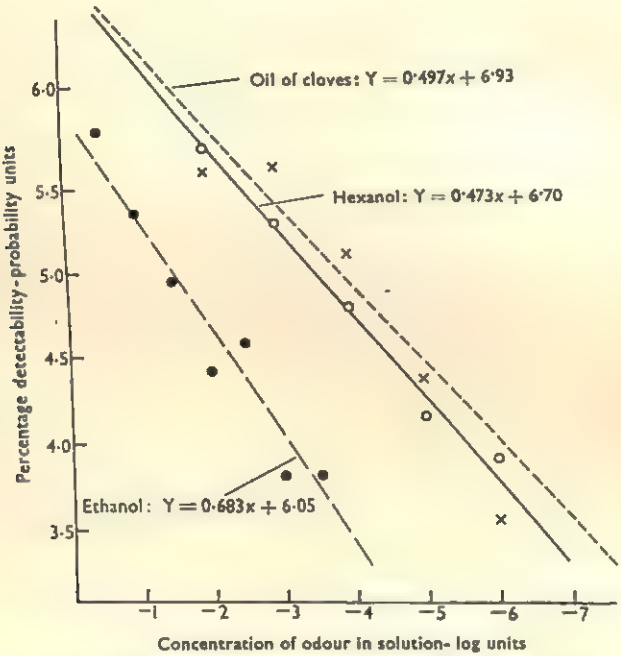


Circular choice apparatus—for description see text.

training, attain scores significantly better than chance in scored blocks of manually conducted trials using strong concentrations of the odour (e.g. oil of cloves). Performance improved as a result of continued trials but 20 trials daily over 3 days was usually sufficient to ensure that only an occasional error was made although rats continued to improve their performance on weaker concentrations over a somewhat longer period.

Figure 3 illustrates the decay in detectability with decreasing concentrations of oil of cloves (based on seven rats), hexanol (seven rats) and ethanol (six rats). The calculated probit regression lines show a very satisfactory fit to the data, and give median threshold values, i.e. the concentrations at which the rat may be expected to respond correctly over 50 per cent. of trials after making allowance for correct responses attributable to chance solution (see Finney, 1952) of 2·9 parts in 100 ml. for ethanol and of 1·7 and 1·3 parts in 100 ml. for hexanol and eugenol respectively.

FIGURE 3



Probit regressions based on the detectability of decreasing concentrations (log to base 10) of ethanol (● and — — —), hexanol (○ and ———) and oil of cloves (× and - - -).

Discussion

It is clear from the observations reported in this paper that the ease with which similar olfactory discriminations can be learned depends to a large extent upon the mode of presentation of the problem. A dominant factor in determining the likelihood of success would seem to be the degree of contiguity of stimulus and reward. Where these were remote, as in the Y-maze, learning did not occur whereas where the odour was directly attached to the drinking spouts, little difficulty was experienced.

The need for such immediate contiguity, which would not seem to be so essential for learning based on stimuli provided by other modalities of sensation, e.g. visual (Lashley, 1930), auditory (Pennington, 1938), provides a satisfactory means of explaining why some earlier workers, but not others, experienced difficulty in establishing olfactory discriminations. Thus in instances where rapid learning has been reported the positive stimulus has either been provided by the odour of the reward itself (Liggett, 1928; Lashley and Sperry, 1943) or, as in the present experiments, applied in close juxtaposition (e.g. French, 1940; Le Magnen and Rapaport, 195

Gruch, 1957). In this connection, the experiments of Swann (1933), who taught rats to burrow through piles of scented sawdust to reach food, suggest that the discrimination can fairly readily (though less readily than in the experience of the workers previously cited) be made at some distance from the reward provided that continuity of stimulus be maintained between the site of stimulus and the reward. On the other hand, in experiments where stimulus and reward have been more widely separated it has proved either difficult or impossible to establish an association between the two. Swann (1933), for example, was unable to teach rats to discriminate odours using a modified Lashley jumping stand while Liggett (1928) was no more successful with a T-maze or a Yerkes discrimination apparatus, except when the test odour used was a trigeminal irritant. The findings of Brown and Ghiselli (1938) who, by applying deterrent and reward in a manner similar to that used in the present experiment with the Y-maze, did succeed in training rats to avoid a pathway identified with the odour of creosote in about 750 trials, suggest that rats trained on the Y-maze might have learned the problem had the phase of training involving the use of electrically-charged doors been pursued further. There is doubt, however, whether creosote may be regarded as stimulating purely olfactory mechanisms (see Allen, 1937) and in any case the number of trials required for training would make the use of this method for determining olfactory thresholds on a statistical basis extremely laborious.

The experimental observations further suggest that, in the rat, the olfactory stimulus plays a relatively small part in orientation, for it is apparent from the behaviour seen in the rectangular choice apparatus after interchange of the positions of the positive and negative drinking tubes, that the animals tend to respond primarily to non-olfactory (presumably visual) cues. When steps are taken to minimize the influence of such cues, either in a rectangular choice box (Eayrs—in preparation) or in the circular choice apparatus already described, rats will more readily react to the olfactory component of the situation. Such observations are consistent with the inability of rats to follow an odorous trial (Liggett, 1928) and of the relatively unimportant part played by olfaction in maze learning by normal, but not by blinded, rats (Honzik, 1936; Casper, 1933).

The technique ultimately developed provides a useful instrument for the study of olfactory acuity and avoids most of the disadvantages, outlined in the Introduction, of methods previously used to study olfaction in animals. Thus training is rapidly and economically accomplished, and the conditions of the experiment can be well controlled. By using water as reward and electric shock as deterrent, and by randomizing the positioning of the odour in relation to drinking tubes and of drinking tubes in relation to other environmental features, the possibility of rats responding to stimuli other than those under test is reduced to minimal proportions. One disadvantage of the apparatus as at present constructed is that, from the point of view of determining olfactory thresholds, the concentrations of odorant molecules reaching the olfactory epithelium must be computed from factors such as their concentration in solution and their rate of diffusion. Only minor modifications are required, however, to enable a more direct measurement to be made by means of air-dilution of the odorant test-substance.

References

ADRIAN, E. D. (1951). Differential sensitivity of the olfactory receptors. *J. Physiol.*, **115**, 42P.

ALLEN, W. F. (1937). Olfactory and trigeminal conditioned reflexes in dogs. *Amer. J. Physiol.*, **118**, 532–40.

BROWN, A. W., and GHISELLI, E. E. (1938). Subcortical mechanisms in learning. IV. Olfactory discrimination. *J. comp. Psychol.*, **26**, 109–20.

CASPER, B. (1933). The normal sensory control of the perfected maze habit of the albino rat. *Psychol. Bull.*, **30**, 702–3.

CLARK, W. E. Le Gros (1956). Observations on the structure and organization of the olfactory receptors in the rabbit. *Yale J. Biol. Med.*, **29**, 83–95.

ELSBERG, C. A., BREWER, E. D., and LEVY, I. (1935). The sense of smell. IV. Concerning conditions which may temporarily alter normal olfactory acuity. *Bull. neurol. Inst. N.Y.*, **4**, 31–4.

FINNEY, D. J. (1952). *Probit Analysis*. London.

FISHER, R. A., and YATES, F. (1948). *Statistical Tables for Biological Agricultural and Medical Research*, Edinburgh.

FRENCH, R. L. (1940). A method and apparatus for demonstrating olfactory sensitivity in the rat. *J. comp. Psychol.*, **29**, 307–13.

GELLERMAN, L. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.*, **42**, 207–8.

GRUCH, W. (1957). Über die Riechfähigkeit bei Wanderratten. *Zool. Jb.*, **67**, 65–80.

GUILLOT, M. (1948). Anosmies partielles et odeurs fondamentales. *C.R. Acad. Sci.*, **226**, 1307–9.

HAND, N. E. (1952). *The Orientation of Animals*. Ph.D. Thesis, University of Birmingham.

HONZIK, C. H. (1936). The sensory basis of maze learning in rats. *Comp. Psychol. Monogr.*, **13**, No. 64.

LASHLEY, K. S. (1930). The mechanism of vision. I. A method for rapid analysis of pattern vision in the rat. *J. genet. Psychol.*, **37**, 453–60.

LASHLEY, K. S., and SPERRY, R. W. (1943). Olfactory discrimination after destruction of the anterior thalamic nuclei. *Amer. J. Physiol.*, **139**, 446–50.

LIGGETT, J. R. (1928). An experimental study of the olfactory sensitivity of the white rat. *Genet. Psychol. Monogr.*, **3**, 1–64.

LE MAGNEN, J. (1950). Nouvelles données sur le phénomène de l'exaltolide. *C.R. Acad. Sci.*, **230**, 1103–5.

LE MAGNEN, J. (1952). Les phénomènes olfacto-sexuels chez le rat blanc. *Arch. Sci. physiol.*, **6**, 295–331.

LE MAGNEN, J., and RAPAPORT, A. (1951). Essai de détermination du rôle de la vitamine A dans le méchanisme de l'olfaction chez le rat blanc. *C.R. Soc. Biol. Paris.*, **145**, 800–3.

MAIER, N. R. F. (1932). The effect of cerebral destruction on reasoning and learning in rats. *J. comp. Neurol.*, **54**, 45–75.

MOULTON, D. J., and EAYRS, J. T. (1960). Study in olfactory acuity. II. Relative detectability of *n*-aliphatic alcohols by the rat. *Quart. J. exp. Psychol.*, **12**, 99–109.

PENNINGTON, L. A. (1938). The function of the brain in auditory localization. IV. Method of training and control experiments. *J. comp. Psychol.*, **25**, 195–211.

PFAFFMAN, C., GOFF, W. R., and BARE, J. K. (1958). An olfactometer for the rat. *Science*, **128**, 1007–8.

SCHNEIDER, R. A., and WOLFF, S. (1955). Olfactory perception thresholds for citral utilizing a new type of olfactorium. *J. appl. Physiol.*, **8**, 337–42.

STONE, C. P. (1941). A simple method for developing olfactory discrimination habits in rats. *J. genet. Psychol.*, **58**, 419–21.

SWANN, H. G. (1933). The function of the brain in olfaction. I. Olfactory discrimination and an apparatus for its test. *J. comp. Psychol.*, **15**, 229–41.

YATES, F. (1936). Incomplete randomized blocks. *Ann. Eugen.*, **7**, 121–40.

STUDIES IN OLFACTORY ACUITY. II.

Relative Detectability of *n*-Aliphatic Alcohols by the Rat

BY

D. G. MOULTON* and J. T. EAYRS†

From the Department of Anatomy, University of Birmingham

The ability of albino and black rats to detect *n*-aliphatic alcohols in the vapour phase has been investigated and probit analysis used to evaluate the results. At median threshold (= 50 per cent. success level) detectability tends to increase by logarithmic increments as the number of carbon atoms in the molecule is increased. However, a similar but reverse relationship occurs between carbon chain length and the gradient of the probit regression lines; and detectability at the 85 per cent. success level, as estimated by interpolation, shows a trend towards oscillation. It is suggested that this finding can resolve the apparent conflict in the literature concerning the pattern of odour intensity in homologous series, and that it may reflect the influence of low water solubility in limiting response to high concentrations of longer chain alcohols.

When expressed as pressures, median threshold values for alcohols in the rat vary directly with saturated vapour pressures; when expressed as thermodynamic activities intermediate and longer chain alcohols appear to be equally stimulating, whilst short chain alcohols show decreasing activities as the series is ascended. In several of the relations considered the position of methanol and dodecanol appears anomalous.

Introduction

Attempts to elucidate the nature of the olfactory process have resulted in an increasing number of studies in which relations between olfactory stimulating efficiency and physico-chemical properties in series of homologous compounds have been considered. In such series chemical and structural differences between successive members are at a minimum, whilst many physical properties alter progressively as the series is ascended. Conditions thus tend to favour an attempt to determine the particular physical properties of a molecule which exert greatest control over its olfactory stimulating efficiency. However, not only is much of the evidence resulting from this approach conflicting, but in many studies statistical design and treatment are inadequate. A further problem is that threshold data relating to species other than man or insects are scarce. The present study, in which the responses of rats to the series of *n*-aliphatic were investigated, was designed to provide further information on these points.

Material and Methods

Apparatus

The circular choice apparatus described in a preceding paper (Eayrs and Moulton, 1960) was used. Throughout these experiments the test odours and control solutions were contained in expendable nylon capsules which were discarded after each trial and replaced by a newly made-up set. The mode of conduct and scoring of the tests was as already described.

Animals

Young rats taken from two stocks initially derived by crossing black and white albino strains of the Norway rat (*Rattus norvegicus*, Berkenhout) were used. Since it has been claimed that olfactory acuity is influenced by the ovarian cycle (Elsberg, Brewer and Levy, 1935; Le Magnen, 1950; Schnedier and Wolff, 1955), only males were tested.

* Present address: Division of Physiology, Florida State University, Tallahassee.
† Henry Head Research Fellow of the Royal Society.

Odorants

The *n*-aliphatic alcohols used, ranging in homologous series from methanol to dodecanol, were supplied by British Drug Houses Ltd., and were redistilled before testing. These, except for a small preliminary group of tests involving the lower members of the series in which distilled water was used as solvent, were first dissolved in 10 per cent. aqueous propylene glycol and subsequently diluted volumetrically with distilled water by successive factors of 10. The 10 per cent. propylene glycol solution used as solvent was similarly diluted to provide a control for the test substance.

Treatment of data

The percentage of correct choices made at each concentration were subjected to probit analysis to determine the median threshold response for each alcohol. The fiducial limits of estimates so derived are not altogether reliable (for discussion see Cheesman and Mayne, 1953) and have accordingly not been used in assessing the results.

Initially concentrations of alcohols eliciting the median threshold response were determined as molar values in solution, but the concentration of the alcohol vapour is clearly a more meaningful unit. In order to estimate this value it is first necessary to know the partial pressure (pt) of the alcohol vapour above the solution. This has been derived from the relation: $pt = ps \times M$, where ps is the saturated vapour pressure of the pure alcohol, and M is the mole fraction (i.e. the proportion of the total molecules of the solution which are alcohol molecules). The quantity ps has been calculated for the temperature of the experiment from tables compiled by Driesbach (1952).

The molar concentration (C) of the alcohol vapour can now be determined by substituting a value for pt and t (the temperature of the experiment), in the equation:

$$C = pt/(22{\cdot}41 \times \frac{t + 273}{273} \times 760) \text{ (cf. Ferguson and Pirie, 1948).}$$

Thermodynamic activities have been obtained from the relationship pt/ps which is the present instance gives a value numerically identical to the mole fraction (cf. Ferguson, 1951).

Results

Results from a preliminary series of experiments in which 3 to 6 brown and albino rats were used are given in Table I. In the case of ethanol, median threshold values were determined using both water and propylene glycol as solvent, and since these values are closely similar they cannot be taken as supporting the view that the choice of solvent influences threshold. It is also clear that detectability tends to increase as the series is ascended. However, since variance at median threshold was high, a second series of trials was undertaken, using a larger number of subjects. This series, in which 27 to 29 black and albino rats were tested on each alcohol, was conducted in two stages: (i) ethanol, butanol, hexanol, octanol, decanol, and

TABLE I

Detectability of *n*-Aliphatic Alcohols in the Vapour Phase by the Rat. Preliminary Series

Alcohol	*Number of presentations*	*Slope of regression line*	*Log molar concentration of solution*	*Log molar concentration of vapour*
Ethanol* ..	864	0·694	−0·343	−4·649
Ethanol ..	576	0·323	−0·507	−4·213
Propanol* ..	864	0·300	−1·324	−5·989
Butanol* ..	864	0·454	−3·052	−7·222
Pentanol ..	864	0·223	−4·269	−9·262
Heptanol ..	384	0·234	−2·031	−8·828
Decanol ..	576	0·289	−4·943	−12·110

* Water used as solvent.

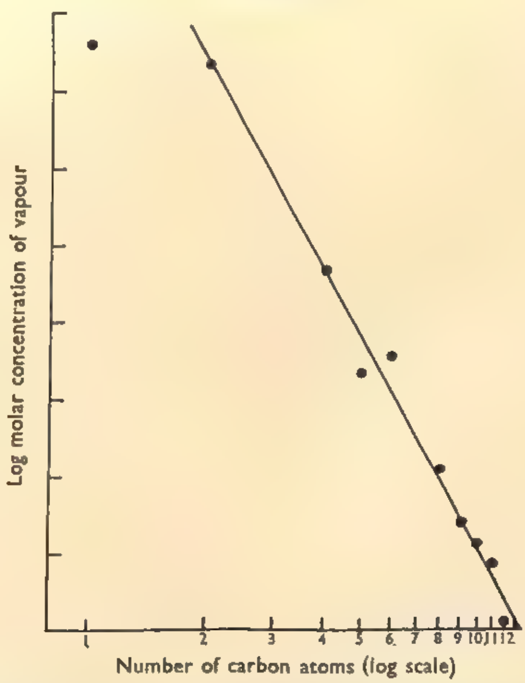
TABLE II

DETECTABILITY OF *n*-ALIPHATIC ALCOHOLS IN THE VAPOUR PHASE BY THE RAT. SECOND SERIES

Alcohol	*Number of presentations*	*Slope of regression line*	*Log molar concentration of solution*	*Log molar concentration of vapour*	*Log activity*
Methanol ..	1512	0·547	−0·9726	−4·4377	−2·1192
Ethanol ..	1624	0·564	−0·9292	−4·6453	−2·0654
Propanol ..	Interpolated from Table I			−5·9890	—
Butanol ..	1624	0·595	−2·7757	−7·3359	−3·9105
Pentanol ..	1512	0·374	−3·2624	−8·6647	−4·3971
Hexanol ..	1624	0·286	−3·0473	−8·4216	−4·1720
Heptanol ..	Interpolated from Table I			−8·8280	—
Octanol ..	1512	0·276	−3·6304	−9·9158	−4·7652
Nonanol ..	1512	0·293	−3·8383	−10·5763	−4·9731
Decanol ..	1512	0·264	−3·8097	−10·8729	−4·9442
Undecanol ..	1512	0·250	−3·7474	−11·1455	−4·8821
Dodecanol ..	1624	0·385	−4·2576	−11·8952	−5·3923

dodecanol; (ii) methanol, pentanol, nonanol, and undecanol using propylene glycol as solvent in each case. The results are given in Table II, in which the molar concentrations for propanol and heptanol, which were tested only in the first series of experiments, have been interpolated in the interest of continuity.

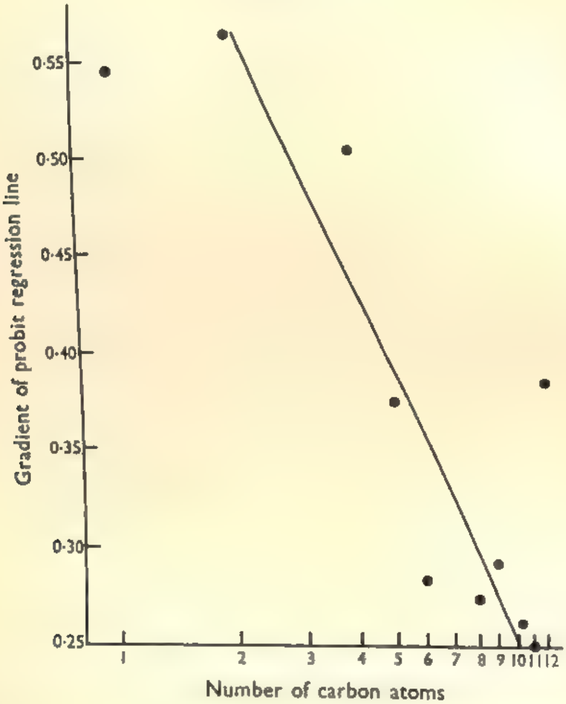
FIGURE 1



Relation of median threshold concentration of homologous series of alcohols, expressed as log molar concentration of vapour, to the number of carbon atoms in the chain. The aberrant value for methanol (one carbon atom) has been excluded from the computation of the regression (see text).

The most striking feature of these results is the tendency for odour intensity to increase by logarithmic increments as the series is ascended. This is shown more clearly in Figure 1, where molar concentrations of alcohol vapour at median threshold are plotted against the number of carbon atoms in the molecule. A linear and logarithmic relation also appears between the slope of the probit regression line and carbon chain length (Fig. 2), indicating that as the series is ascended there is a progressive reduction in the amount by which the response increases with each unit change in concentration. For reasons discussed below methanol has been excluded in computing the regressions of chain length on molar concentration and on slope

FIGURE 2



Relation of slope of probit regression lines for the detectability of alcohols forming a homologous series to the number of carbon atoms in the chain. The aberrant values for methanol and dodecanol have been excluded from the computations (see discussion).

(Figs. 1 and 2). In Figure 2 the position of dodecanol is clearly aberrant in comparison with the remaining members of the series and for this reason it also has been excluded from the computation of the regression line.

One implication of these findings is that the pattern of detectability found is dependent on the criterion of response chosen. This is immediately apparent if the relative stimulating efficiency of the alcohols at the 50 per cent. success level is compared with the pattern derived, by substitution in the probit regression equations from the 85 per cent. success level (Figs. 1 and 2). At these higher concentrations there is a clear trend towards oscillation in response, alcohols with 4 to 11 carbon atoms occupying a concentration range of 2·3 log units as compared with 3·8 log units at the 50 per cent. success level. This trend continues until response to undiluted alcohols is being estimated.

Discussion

Odour intensity in homologous series

Among those who have also reported that the detectability of alcohols increases as the series is ascended are Cook (1926), Dethier and Yost (1952), and Hughes (1957), using flies, and Passy (1892), Skramlik (1948) and Kruger, Feldzamen and Miles (1955) in man. (However, the claims of the last workers refer only to the longer chain alcohols and only after corrections for vapour pressure are applied.) A similar relation to chain length is seen in human olfactory thresholds for alkyl mercaptans and in the amplitudes of the slow potentials evoked by ketones and aldehydes in the olfactory mucosa of the frog (Allison and Katz, 1919; Ottoson, 1958).

In the present study, in data given by Passy (1892) and in the responses of blowflies to alcohols, aldehydes and fatty acids (Dethier and Yost, 1952; Dethier, 1954*b*; Hughes, 1957) this relationship is clearly logarithmic. There is also some evidence from work on the dog (Ashton, Eayrs and Moulton, 1957), which has since been repeated, that the increasing detectability of fatty acids as the series is ascended is influenced by some factor which likewise alters logarithmically in value. Thus the generalization that, in several series, odour intensity rises as chain length expands has considerable support. It also accords with the action of certain of these series on a number of other biological systems. For example, the efficiency of alcohols in accelerating haemolysis (Ponder and Hyman, 1939), in producing a variety of narcotic effects (see Brink and Posternack, 1948), in stimulating the tarsal chemoreceptors of blowflies and the taste receptors of man (Dethier and Chadwick, 1947; see Dethier, 1956) increases by logarithmic increments as chain length is increased.

In contrast, however, the olfactory response to the homologous series of the fatty acids show one or more oscillations in detectability as the series is ascended—the first peak of stimulating efficiency being related to members with 3 to 5 carbon atoms, e.g. the studies of Passy (1893), Backman (1917*a*) and Skramlik (1948) in man, Neuhaus (1953) on the dog, Ottoson (1958) on the frog, and Schwarz (1955) on the honey bee. A similar oscillation has been found in an alcohol, a paraffin and a chloroparaffin series (Mullins, 1955*a*) and Ottoson (1958) reports that in an alcohol series the peak of maximum activity is related to the members with 7 to 8 carbon atoms.

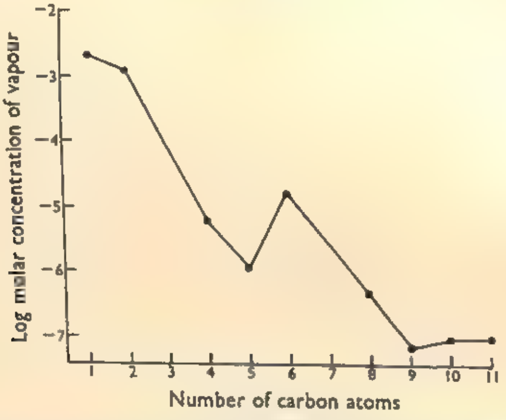
It thus appears that there exist two strongly conflicting sets of data concerning the relationship between intensity of odour and position in a homologous series; and although differences in the species and techniques used by various workers doubtless complicate the problem they seem unlikely to offer a complete solution since different patterns have been reported in the same species (e.g. Mullins, 1955*a*, Passy, 1892) and by the same worker (Ottoson, 1958). In the following paragraphs, therefore, an attempt is made to reconcile these data in the light of the current findings.

Significance of relationship between intensity of response and concentration of vapour

(*a*) *As a means of reconciling divergences of evidence concerning the pattern of response.* The conclusion drawn from a comparison of Figures 1 and 2 that the pattern of detectability may vary with the criterion of response chosen offers a possible explanation of how these conflicting claims have arisen. Thus the lower the success score that is chosen to represent threshold, the greater the probability that a linear relationship between stimulating efficiency and chain length will be found. However, if it is assumed that similar concentrations of alcohol vapours above the olfactory organ of

different subjects tend to produce similar patterns of response, variations in olfactory sensitivity between species could also have importance. For example, a criterion for success in the region of 85 per cent. is required to reveal an oscillatory pattern in the responses of rats to alcohols (Fig. 3), whereas in man Mullins (1955*a*) found an oscillation in detectability using what is clearly a much lower success value (i.e. the level at which subjects "found difficulty in recognizing every stimulus"). The present results suggest that the concentration of alcohol vapour required to elicit a low success level in man would elicit a much higher success level in rats, and thus, if the assumption holds that similar concentrations yield similar patterns of response, an explanation of such apparent anomalies is available.

FIGURE 3



Relation of alcohol concentration, corresponding to a 85 per cent. criterion of success, to the number of carbon atoms in the molecule. Concentrations were estimated by substitution in the equation for the probit regression for each alcohol.

Doubtless differences in techniques, methods of analysis, purity of test compounds, amount of mucus in the nasal cavity and similar variables may also contribute to the complexity of the picture. Kruger, Feldzamen and Miles (1955), for example, were unable to find any change in the pattern of intensity over "a wide range of dilutions" of alcohols (molar concentrations were not specified). But these workers were using a subjective iso-intensity odour scale and it is difficult to know what relation this criterion of response bears to that used by other workers.

However, the possible cause of this variation in response pattern merits consideration since it throws light on certain aspects of the olfactory process.

(*b*) *As a possible cause of the transformation of the pattern of response.* Oscillation in biological activity within an homologous series of compounds is generally considered to be the result of interaction between two opposing factors, the first tending to cause an increase in activity as the series is ascended, and the second having the opposite effect. According to Sexton (1953), a common explanation of the second phase is that at a certain point in the series the concentration of a compound needed to produce a given effect exceeds its water solubility. At this stage the addition of further quantities of the compound will not increase activity.

There are grounds for believing that an effect of this kind could largely account for the pattern of response seen in Figure 3. Thus the progressively increasing detectability of the short chain alcohols is associated with complete miscibility in

water, whilst in the longer chain alcohols the decay in stimulating efficiency corresponds with diminishing water solubility. If this explanation is correct the transformation of response pattern from the oscillating to the linear form (cf. Figs. 1 and 3) could be related to the declining importance of water solubility as a limiting factor and thus to the emergence of whatever property controls activity in the short chain alcohols as the dominant activity throughout the series. The logarithmic relation between chain length and slope in alcohols of limited solubility accords with this view since water solubility also decreases by logarithmic increments as chain length is increased.

However, response increments for alcohols of 5 to 12 carbon atoms not only fail to continue their decline at high concentrations, but show a marked increase in magnitude. Thus if the theory is correct it is necessary to postulate the intervention of a further factor reinforcing response at high concentrations.

In the case of the apparently similar effect which occurs in the responses of the honey bee to the vapours of acetic acid and phenol Jones (1952) has suggested that at high concentrations the vapours act only as irritants whereas at low concentrations the olfactory sense alone is effective. But although all objective evidence available supports the view that in both insects and mammals irritants irritate only in high concentrations (see Parker and Stabler, 1913; Dethier, 1953; Neuhaus, 1953), it is premature to consider whether the present anomalies are due to the intervention to trigeminal response or to some other factor such as direct chemical action or the excitation of different types of olfactory or other receptors. It is clear that much more work will be required to elucidate the nature of these effects.

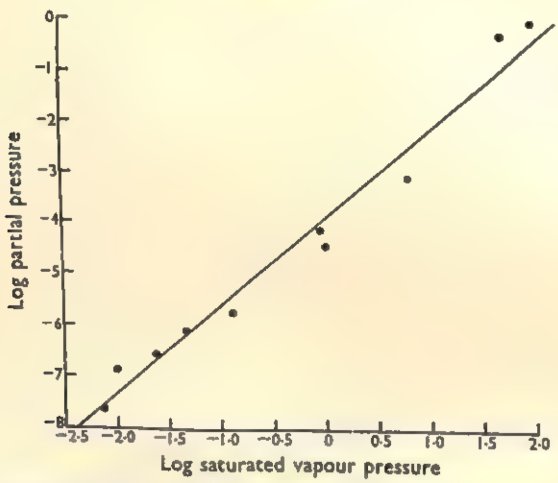
Implications for an understanding of the nature of the olfactory process

If, as seems highly probable, the olfactory hairs or corpuscles (cf. Hopkins, 1926; Clark, 1957) are covered by a thin film of mucus or adsorbed moisture, reduced water solubility may restrict the stimulating efficiency of the alcohols by limiting their access to the receptor surface and may in this way be responsible for the oscillatory pattern of detectability. If this were so such a pattern could hardly give a true indication of the basic sensitivity of the receptors to alcohols. But in any case the pattern of increase in response to alcohols as the series is ascended (Fig. 1) seems here more likely to provide insight into the nature of the olfactory process since at median threshold anomalies of the type occurring at higher and lower levels of success have less influence. At this level the odour intensity of the alcohols is clearly determined by some property which tends to alter by logarithmic increments as chain length is increased.

Ferguson (1939) and others have pointed out that several physical properties, including oil-water partition coefficients, surface activity and vapour pressure show a similar relationship to chain length within an homologous series. Thus when the median thresholds for alcohols given in Table I are expressed as pressures and plotted against their saturated vapour pressures (Fig. 4) the relation is clearly linear. This type of relation is characteristic of narcotic and toxic phenomena in which it is assumed that an equilibrium exists between the external phase and the internal biophase. Ferguson (1939) has argued that when such an equilibrium exists the thermodynamic activity of the test compound will have the same numerical value in all phases. Thus the activity of the substance at the site of biological action can be derived from a knowledge of its concentration outside the organism. Thermodynamic activity is defined as partial free molal energy referred to a standard state and is numerically equal to the relative saturation of the vapour of the compound (i.e. pt/ps, see Ferguson and Pirie, 1948). (For a fuller discussion of the rationale

behind the application of analysis in terms of thermodynamic activities, see Ferguson (1939, 1951), Ferguson and Pirie (1948), Brink and Posternack (1948) and Dethier (1954*a*).) Applying this form of analysis to data on the olfactory responses of blowflies, Dethier and Yost (1952) and Dethier (1954*b*) found that thresholds for *n*-aliphatic alcohols and aldehydes of intermediate chain length were approximately equally stimulating and they therefore concluded that in these series olfaction may involve an equilibrium process. Similarly Ottoson (1958), in a study of the slow potentials evoked in the olfactory mucosa of the frog, has reported that alcohols of intermediate chain length stimulate at equal thermodynamic activities.

FIGURE 4



Relation of median threshold, expressed as partial vapour pressure, to saturated vapour pressure.

It seems pertinent therefore to examine the pattern of median threshold values, reported above, in terms of activity units (Table I). In Figure 5 these activities are plotted against chain length and the resulting pattern suggests that, although they show an overall tendency to decrease with increasing chain length, the alcohols tested (with the exception of methanol and dodecanol) tend to fall into two groups:

(1) Members with 2 to 4 carbon atoms which are detected at decreasing activities as the series is ascended.

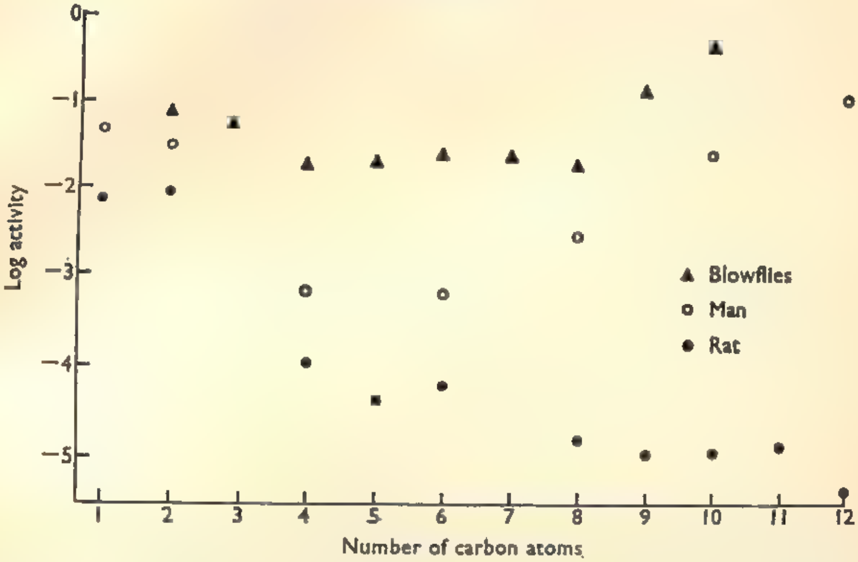
(2) Members with 5 to 11 carbon atoms which are approximately equally stimulating.

If the behaviour of the longer chain alcohols is excluded such a conclusion is essentially in agreement with the findings of Dethier and Yost (1952), Dethier (1954*b*) and Ottoson (1958) whilst an analysis of data given by Ashton, Eayrs and Moulton (1957) suggests a similar division in a fatty acid series. However, although the trend shown by the lower alcohols accords with all available evidence some studies show a different pattern in remaining members. Thus, according to an analysis made by Dethier (1954*a*) of scattered data on olfactory response for man and for flies (Cook, 1926) in a number of homologous series, a decrease in activities occurs as the series is ascended. Further, the oscillation which Mullins (1955*a*) reports

in odour intensity in three series, is also shown when thresholds are expressed as activities (Mullins; 1955*b*; see also Fig. 5) and a similar pattern appears in more meagre data on human olfactory thresholds for alcohols (Gavaudan, Brebion, Poussel and Shutzenberger, 1948).

Further information is clearly needed, but it may be tentatively concluded that whilst olfaction in the rat for medium and long chain alcohols may involve an equilibrium process, olfaction for shorter chain alcohols may not. Ottoson (1958) has suggested that solubility properties may account for a similar break which he reports in this series.

FIGURE 5



Relation of median threshold concentrations, expressed as thermodynamic activities, to chain length in a series of *n*-aliphatic alcohols. Olfactory responses of rats are compared with those of blowflies (Dethier and Yost, 1952), and man (Mullins, 1955*b*).

A further indication of the complexity of factors which may influence the detectability of alcohol vapour is seen in the frequent failure of methanol and dodecanol to conform to the trends shown by the remaining alcohols (Figs. 1, 2 and 5). Other workers have noted, or their data suggest, the greater stimulating efficiency of methanol than its position in the series would seem to warrant (Passy, 1892; Dethier and Yost, 1952) and such anomalies are also a feature of toxicity data (see Passy, 1892; Ostergren, 1944; Ferguson and Pirie, 1948). The aberrant behaviour of the high alcohols in their narcotic effects has been noted by Brink and Posternack (1948).

The present study tends to support the suggestions of Beckman (1917*b*), Ottoson (1958) and others that the solubility characteristics of compounds may play an important part in determining their olfactory stimulating efficiency, and indicates that the mechanisms of olfaction may not differ greatly in the rat, in the frog (Ottoson, 1958), and in blowflies (Dethier and Yost, 1952). Finally, the similarity between the present findings and those of Ottoson (1958) suggests that olfactory data obtained by behavioural techniques can closely reflect the patterns of comparable data obtained by use of electrophysiological methods.

REFERENCES

ALLISON, V. C., and KATZ, J. H. (1919). An investigation of stenches and odours for industrial purposes. *J. indust. Eng. Chem.*, **11**, 336–8.

ASHTON, E. H., EAYRS, J. T., and MOULTON, D. G. (1957). Olfactory acuity in the dog. *Nature*, **179**, 1069–70.

BACKMAN, E. L. (1917*a*). Upsala Läkareför. *Förhandl.*, **22**, 329 (cited by Hoffman, F. B. (1926) in *Handbuch der Normalen und Pathologischen Physiologie*, **11**, 253–399).

BACKMAN, E. L. (1917*b*). Note sur la puissance des odeurs et leur solubilité dans l'eau et dans l'huile. *J. Physiol. Path. gen.*, **17**, 1–4.

BRINK, F., and POSTERNACK, J. M. (1948). Thermodynamic analysis of the effectiveness of narcotics. *J. cell. comp. Physiol.*, **32**, 211–33.

CHEESMAN, G. H., and MAYNE, S. (1953). The influence of adaptation on absolute threshold measurements for olfactory stimuli. *Quart. J. exp. Psychol.*, **5**, 22–30.

CLARK, W. E., Le Gros (1957). Inquiries into the anatomical basis of olfactory discrimination. *Proc. Roy. Soc. B.*, **146**, 299–319.

COOK, W. C. (1926). The effectiveness of certain paraffin derivatives in attracting flies. *J. agric. Res.*, **32**, 347–58.

DETHIER, V. G. (1953). Chemoreception. In Roeder, K. D. (editor). *Insect Physiology*. London. 547–76.

DETHIER, V. G. (1954*a*). The physiology of olfaction in insects. *Ann. N.Y. Acad. Sci.*, **58**, 139–57.

DETHIER, V. G. (1954*b*). Olfactory responses of blowflies to aliphatic aldehydes. *J. gen. Physiol.*, **37**, 743–51.

DETHIER, V. G. (1956). Chemoreceptor mechanisms. In Grenell, R. G., and Mullins, L. J. (editors). *Molecular Structure and Functional Activity in Nerve Cells*. Washington, 1–33.

DETHIER, V. G., and CHADWICK, L. E. (1947). Rejection thresholds of the blowfly for a series of aliphatic alcohols. *J. gen. Physiol.*, **37**, 823–39.

DETHIER, V. G., and YOST, M. G. (1952). Olfactory stimulation of blowflies by homologous alcohols. *J. gen. Physiol.*, **35**, 823–39.

DRIESBACH, R. R. (1952). *Pressure—Volume—Temperature Relationships of Organic Compounds*. Sandusky.

EAYRS, J. T., and MOULTON, D. G. (1960). Studies in olfactory acuity. I. Measurement of olfactory thresholds in the rat. *Quart. J. exp. Psychol.* **12**, 90–98.

ELSBERG, C. A., BREWER, E. D., and LEVY, I. (1935). The sense of smell. IV. Concerning conditions which may temporarily alter normal olfactory acuity. *Bull. Neurol. Inst. N.Y.*, **4**, 31–4.

FERGUSON, J. (1939). The use of chemical potentials as indices of toxicity. *Proc. Roy. Soc. B.*, **127**, 387–404.

FERGUSON, J. (1951). Relations between thermodynamic indices of narcotic potency and the molecular structure of narcotics. In Méchanisme de la Narcose. *Coll. int. Cent. nat. Rech. Sci.*, **26**, 25–39.

FERGUSON, J., and PIRIE, H. (1948). The toxicity of vapours to the grain weevil. *Ann. appl. Biol.*, **35**, 532–50.

GAVAUDAN, P., POUSSEL, H., BREBION, G., and SHUTZENBERGER, M. P. (1948). L'étude des conditions thermodynamiques de l'excitation olfactive et les théories de l'olfaction. *C.R. Acad. Sci.*, **226**, 1395–6.

HOPKINS, A. E. (1926). The olfactory receptors in vertebrates. *J. comp. Neurol.*, **41**, 253–89.

HUGHES, J. C. (1957). Olfactory stimulation of tsetse flies and blowflies. *Bull. ent. Res.*, **48**, 561–79.

JONES, G. D. C. (1952). The responses of the honey bee to repellant chemicals. *J. exp. Biol.*, **29**, 372–86.

KRUGER, L., FELDZAMEN, A. N., and MILES, W. R. (1955). Comparative olfactory intensities of the aliphatic alcohols in man. *Amer. J. Psychol.*, **68**, 386–95.

LE MAGNEN, J. (1950). Nouvelles données sur le phénomène de l'exaltolide. *C.R. Acad. Sci.*, **230**, 1103–5.

MULLINS, L. J. (1955*a*). Olfactory thresholds for some homologous series of compounds. *Fed. Proc.*, **14**, 105–6.

MULLINS, L. J. (1955*b*). Olfaction. *Ann. N.Y. Acad. Sci.*, **62**, 247–56.

NEUHAUS, W. (1953). Über die Riechschärfe des Hundes für Fettsäuren. *Z. vergl. Physiol.*, **35**, 527–52.

OSTERGREN, G. (1944). Colchicine mitosis, chromosome contraction, narcosis, and protein chain folding. *Hereditas*, **30**, 429–67.

OTTOSON, D. (1958). Studies on the relationship between olfactory stimulating effectiveness and physico-chemical properties of odorous compounds. *Acta physiol. scand.*, **43**, 167–81.

PARKER, G. H., and STABLER, E. M. (1913). On certain distinctions between taste and smell. *Amer. J. Physiol.*, **32**, 230–40.

PASSY, M. J. (1892). L'odeur dans la série des alcohols. *C.R. Soc. biol.*, **4**, 447–9.

PASSY, M. J. (1893). L'odeur dans la série grasse. *C.R. Soc. biol.*, **5**, 479–81.

PONDER, E., and HYMAN, C. (1939). Acceleration of haemolysis in relation to chemical structure. 2. The straight chain alcohols. *Proc. Soc. exp. Biol.*, **42**, 320–2.

SCHWARZ, R. (1955). Über die Riechschärfe der Honigbiene. *Z. vergl. Physiol.*, **37**, 180–210.

SEXTON, W. A. (1953). *Chemical Constitution and Biological Activity*. London.

SCHNEIDER, R. A., and WOLFF, S. (1955). Olfactory perception thresholds for citral utilizing a new type of olfactorium. *J. appl. Physiol.*, **8**, 337–42.

SKRAMLIK, E. von (1948). Über die zur minimalen Erregung des menschlichen Geruchs- bzw. Geschmackssinnes notwendigen Molekülmengen. *Pflüg. Arch. ges. Physiol.*, **249**, 702–16.

SHORTER ARTICLES AND NOTES

TWO MORE VISUAL THEOREMS

BY

G. S. BRINDLEY

From the Physiological Laboratory, Cambridge

The basis of the widely held opinion that the spectral sensitivities of the mechanisms determining trichromacy must be linearly related to the spectral mixture functions is examined, and a proof from simple assumptions is stated.

It is proved also that, on the van der Velden hypothesis of independent sensitive units, a spatial or temporal summation law of the form $AI^n = c$ implies a frequency-of-seeing curve of the form $Q = 1 - e^{-KI^n}$ (Q = probability of detecting a stimulus of size A and intensity I; n, c and K are constants).

Introduction

This theoretical paper is a sequel to one already published in this *Journal* (Brindley, 1957), and presents and proves two more theorems with applications in visual physiology and psychology. These are numbered 3 and 4, following the two of the former paper.

Theorem 3 is of respectable antiquity. First propounded and used by König and Dieterici (1886), it seems to have been regarded by them as self-evident, and I cannot discover that any subsequent writer has published a proof of it, though many have accepted and used it. The chief value of now devising a proof is that the assumptions which underlie its general acceptance are thereby clarified, and we are the better able to judge whether in given circumstances it is likely to be exactly true, approximately true, or altogether false.

Theorem 4 is new. Its application is to the testing of the extent of validity of the important hypothesis, first clearly set out by van der Velden (1944), that the spatial or temporal segments of a large or long visual stimulus are detected *independently* by the retina and visual pathway. This application is made in detail elsewhere (Brindley, 1960, Chapter 6).

Definition

The reciprocal of the amount of light of wavelength λ required to produce a given constant response in a mechanism is called the *spectral sensitivity* of that mechanism for the wavelength λ. The "amount" must be measured in units which remain meaningful despite a change of wavelength; customarily, units either of energy or of number of quanta (per second per steradian, etc.) are used, the former being experimentally more direct, the latter better for correlation with photochemical properties.

If the substitution postulate holds for a mechanism, then its spectral sensitivity will be invariant (except for a factor independent of wavelength) when the response-criterion at which it is measured is changed. This invariance, however, is no essential part of the concept of a spectral sensitivity.

Statement of the Theorems

Theorem 3

In trichromatic vision, the spectral sensitivities of the three mechanisms determining the trichromacy must be linear functions, with coefficients independent of test wavelength, of the amounts of any three specified primaries required to match a constant amount of a test light of variable wavelength.

Theorem 4

If the relation between the intensity (I) and temporal or spatial size (A) of a visual stimulus required to attain a constant probability of seeing is of the form $AI^n = c$, where n and c are constants, and if, over the range of sizes considered, all stimuli are divisible, temporally or spatially, into a large number of similar constituents the detection of which is independent, then at any given size of stimulus the probability of seeing (Q) must be related to the intensity by $Q = 1 - e^{-KI^n}$, where K is a positive constant.

Assumptions Required in the Proof of Theorem 3

We shall assume the postulate of unidimensional response and the substitution postulate (Brindley, 1957). We shall also use Case 1 of Theorem 2. This does not constitute a further assumption, since it is deducible from the first two and the experimental law of uniqueness of colour matches.

Proof of Theorem 3

Let r_λ, g_λ, and b_λ be the amounts of the primaries, r, g and b required to match unit amount of light of wavelength λ. One of these amounts must be negative; we may suppose that it is b_λ.

Let ρ_r, ρ_g, ρ_b and ρ_λ be the spectral sensitivities of the red-sensitive mechanism to r, g, b and λ respectively, and let λ be such that $\rho_\lambda \neq 0$. Then this mechanism is affected to the same extent

by amount $\rho_r r_\lambda/\rho_\lambda$ of wavelength λ
as by amount r_λ of primary r (Subst. post.),

and by amount $\rho_g g_\lambda/\rho_\lambda$ of wavelength λ
as by amount g_λ of primary g (Subst. post.),

and by amount $\rho_b b_\lambda/\rho_\lambda$ of wavelength λ
as by amount b_λ of primary b (Subst. post.),

and by amounts r_λ of primary r plus g_λ of primary g as by amounts $-b_\lambda$ of primary b plus 1 of wavelength λ (Case 1 of Theorem 2, since these match),

and therefore by amount $\rho_r r_\lambda/\rho_\lambda + \rho_g g_\lambda/\rho_\lambda$ of wavelength λ as by amount $1 - \rho_b b_\lambda/\rho_\lambda$ of wavelength λ (Subst. post.).

Therefore $\rho_r r_\lambda/\rho_\lambda + \rho_g g_\lambda/\rho_\lambda = 1 - \rho_\lambda b_b/\rho_\lambda$, for if they were not equal, application of the substitution postulate would prove $\rho_\lambda = 0$, contrary to the assumption made above.

Therefore $\rho_\lambda = \rho_r r_\lambda + \rho_g g_\lambda + \rho_b b_\lambda$.

The theorem is thus proved, since ρ_r, ρ_g and ρ_b are independent of λ.

This proof does not involve the concept of extra-spectral "fundamental stimuli" (Wright, 1934, and other writers), with which it is customary to introduce the spectral sensitivities of the mechanisms determining trichromacy. Proofs can be devised that start from "fundamental stimuli"; but those that I have been able to invent involve the same assumptions as are made above, and also others more difficult to justify.

Of the assumption needed for the proof, the law of uniqueness of colour matches is experimental, and the postulate of unidimensional response is necessarily involved in the concept of a spectral sensitivity. Only the substitution postulate remains to be doubted. Is Grassmann's third law sufficient evidence for its truth? Logically it is not; Grassmann's third law can be deduced from the substitution postulate (given Theorems 1 and 2), but not the converse. However, by scientific standards the argument from Grassmann's law to the substitution postulate is a good one; the substitution postulate is equivalent to a fairly small extrapolation from the verifiable instances of the law, and its credibility is enhanced by its being implied by the hypothesis that the three mechanisms determining trichromacy are receptive pigments, for which there is much independent evidence (see Brindley, 1960, Chapter 7).

Proof of Theorem 4

If the number of independently detected constituents of a stimulus is sufficiently large, boundaries can be neglected, and the number be assumed proportional to the size.

Let the number of independently detected constituents in a stimulus of size A be rA, let the probability of detection of each one of them be P, and let the probability of detection of the whole stimulus be Q.

Then $1 - Q = (1 - P)^{rA}$,

therefore $ln\ (1 - Q) = rA\ ln\ (1 - P)$.

Let Q_o be the value of Q for which the law $AI^n = c$ has been experimentally established.

$$\text{Then } ln\ (1 - Q_o) = (rc/I^n) ln\ (1 - P)$$
$$= (c/AI^n) ln\ (1 - Q),$$
$$\text{therefore } ln\ (1 - Q) = (A/c) ln\ (1 - Q_o).I^n$$
$$= -(A/c) ln \frac{1}{1 - Q_o} .I^n$$

Therefore $Q = 1 - e^{-KI^n}$, where K is the positive constant $(A/c)\ ln\ \dfrac{1}{1 - Q_o}$

The frequency-of-seeing curve (i.e. the curve of Q plotted against log I) is thus determined by the empirically known law of spatial (or temporal) summation and the hypothesis that the detection of constituents is independent. The *shape* of the curve is independent also of the size of stimulus with which it is tested (provided that this is within the range over which $AI^n = c$), and of the probability-criterion Q_o at which $AI^n = c$ was established, for varying A or Q_o merely shifts the curve along the axis of log I.

References

Brindley, G. S. (1957). Two theorems in colour vision. *Quart. J. exp. Psychol.*, **9**, 101–4.

Brindley, G. S. (1960). *Physiology of the Retina and Visual Pathway.* London.

König, S., and Dieterici, C. (1886). Die Grundempfindungen und ihre Intensitätsverteilung im Spektrum. *S.B. Akad. Wiss. Berlin*, 805–29.

Van der Velden, H. A. (1944). Over het aantal lichtquanta dat nodig is voor een lichtprikkel bij het menselijk oog. *Physica.*, **11**, 179–89.

Wright, W. D. (1934). The measurement and analysis of colour adaptation phenomena. *Proc. Roy. Soc.*, **115B**, 49–87.

A NOTE ON SOME UNUSUAL HANDEDNESS PATTERNS*

BY

S. M. GILLIES, D. A. MACSWEENEY and O. L. ZANGWILL

From the Psychological Laboratory, University of Cambridge

It is widely held that the "ambidextrous" individual is almost always a natively left-handed person, trained to use the right hand but still naturally skilful with the left (Wile, 1934; Brain, 1945; Burt, 1950). Whatever the truth of this view, it is at all events clear that left-handed individuals who have been obliged to write with the right hand commonly show a greater variety of hand preferences than similar individuals who have been permitted to write with the left (Hildreth, 1949; Humphrey, 1951). At the same time, it is noteworthy that right-hand preference for certain tasks is not uncommon even among sinistrals who have been allowed—or even encouraged—to persist in their left-handed writing (Humphrey, 1951). For instance, no fewer than ten of the 35 left-handed writers studied by Humphrey (1951) were stated to prefer the right hand for throwing, for unimanual games, or for both these activities. It is evident, therefore, that enforced shifting to the use of the right hand in writing cannot be the sole cause of "mixed" hand preference. In some cases, at least, a somewhat different explanation of "ambidexterity" must be sought.

In an enquiry modelled upon that of Humphrey (1951), ten subjects have been found who are left-handed writers but preferentially right-handed for a variety of of other activities. These represent roughly 15 per cent. of the total group of left-handed subjects so far studied. Of the ten subjects, five consider themselves to be left-handed, four ambidextrous, and one described herself as "split" in her hand preferences. The object of this note is to give a short account of the lateral preference patterns in this atypical group.

Procedure and Subjects

The questionnaire devised by Humphrey (1951) was used in this study. It consists of 20 questions designed to assess hand preference in a number of skilled acts which are common to most educated people. No account is taken of possible differences in strength or skill as between the two hands. From the answers given, an approximate "sinistrality index" (S.I.) is calculated for each subject from the formula $(L + \frac{1}{2}E)/N$ where L and E refer to the number of questions answered "left" and "either" respectively and N is the total number of questions. An S.I. of 1·0 would thus indicate consistent left preference and an S.I. of 0·0 consistent right preference. An index of 0·5 indicates virtual "ambilaterality." We have also followed Humphrey's procedure in adopting an arbitrary "handedness classification" in terms of the preferred hand in writing, throwing and unimanual games such as tennis or squash. Thus L – R – R (which was the pattern displayed in most of our subjects) indicates a left-handed writer who prefers to throw a ball and use a tennis racket with his right hand.

The subjects, all of whom were young University graduates or undergraduates were asked some additional questions concerning ocular and foot preference and certain other matters presumed to bear on laterality, e.g. relative skill of the two hands, tendency to mirror writing, and proneness to confuse left and right. Information regarding familial handedness was also sought. Finally the time taken by every subject to sign his name and write a short text to dictation with either hand was recorded.

* Based on a short communication to the Experimental Psychology Society at a meeting held in Cambridge on 15th July, 1959.

Results

The principal findings are shown in Table I. It will be seen first, that the S.I.s are distinctly low, indicating an appreciable admixture of right hand preferences in all subjects. (Indeed, in the case of the first three subjects, writing and drawing are the only major skills for which the left hand is consistently preferred.) Secondly, the incidence of right foot and eye preference is distinctly high, six subjects being both right-footed and right-eyed and only one being both left-footed and left-eyed.

TABLE I

Lateral Preference Patterns in 10 Left-handed Writers

Subject	*Sex*	*Age*	*S.I.*	*Handedness classification*	*Foot preference*	*Eye preference*	*Familial sinistrality*
1. V.P.B.	F	21	0·15	L – R – R	E	R	Maternal grandmother; brother
2. C.D.M.	M	21	0·15	L – R – R	R	L	Maternal grandmother
3. A.B.	F	21	0·18	L – R – R	R	R	None ascertained
4. C.L.	F	24	0·35	L – R – R	L	L	Two first cousins
5. P.A.J.	M	23	0·38	L – R – R	? R	L	Maternal grandmother; mother; brother; nephew
6. G.W.	F	35	0·40	L – R – R	R	R	Paternal grandfather
7. J.W.A.	M	21	0·45	E – L – R	R	R	Mother
8. M.G.G.	M	20	0·53	L – R – R*	R	R	One maternal uncle
9. D.M.A.S.	M	21	0·58	L – R – R*	R	R	None ascertained
10. C.M.	M	22	0·65	L – R – R	R	R	None ascertained

S.I. denotes sinistrality index.

"Handedness classification" summarizes manual preference with regard to writing, throwing and unimanual games (Humphrey, 1951).

R* indicates right-handed for games, with the exception of table tennis.

E denotes lack of consistent unilateral preference.

Thirdly, there is evidence of familial sinistrality in seven cases, though in only two is a parent reported as left-handed. In three cases no evidence of familial sinistrality could be ascertained.

It is also of interest that three subjects (Nos. 1, 2 and 7) confessed to noteworthy lack of skill in either hand—suggestive of the classical "ambilaevous" condition first described by Galen—and two (Nos. 3 and 7) described some difficulty in discriminating right and left, especially when giving directions as to routes. Seven subjects had some facility in mirror writing.

On the *writing tests*, speed and quality of writing with the non-preferred (right) hand were, in general, but little inferior to performance with the preferred hand, testifying

to appreciable ambidexterity. (In one case (No. 7) writing performance with the two hands was virtually identical, but this subject had practised writing with his non-preferred hand since the age of 11. No other subject reported having practised writing with the right hand.)

Discussion

These unusual preference patterns might be explained in various ways. First, it might be argued that some, at least, of our subjects were naturally right-handed, their left-hand writing being attributable to some incidental psychological perversity. In this connection, Scheidemann and Colyer (1931) have suggested that isolated left-handed writing in an otherwise right-handed child may derive simply from imitation, conscious or unconscious, of a left-handed parent, teacher or admired school fellow. In such cases, these authors claim, reversal of the writing hand is as a rule both easy and successful. As applied to our present cases, this interpretation might gain support from the relatively good writing performance of the subjects with the non-preferred hand, even in the absence of special practice. At the same time, no direct evidence was obtained which suggested that left-handed writing habits had in fact been acquired in this way. No subject was able to recall an early identification with a left-handed writer and should any such have occurred it must almost certainly have remained unconscious. There are also certain other considerations which militate against this explanation as generally applicable to the left-handed writer (cf. Burt, 1950).

A more plausible explanation might perhaps be given in terms of training and social convention. Thus it might be supposed that our subjects were natively left-handed and that their tendency to write with the left hand had been fostered by a permissive attitude at home and at school. At the same time, this permissiveness may not have extended in the same way to other types of unimanual activity normally carried out in a social context. For instance, almost all left-handed children come to adopt right-handed table manners and the urge to conform may likewise govern the adoption of right hand preferences in the use of tools and in certain games. None the less, it is difficult to see why social pressures of this kind should affect some left-handed writers so much more than others, though it is possible that considerations of temperament are involved to some extent (Blau, 1946). It is also possible that some left-handed people have undertaken a measure of self-training in the use of the right hand, perhaps in the belief that ambidexterity confers a real advantage upon its possessor. None the less, it might seem rather doubtful whether training and convention alone are sufficient to explain the high frequency of right hand preferences in the present group.

A third line of explanation—and one to which we ourselves incline—involves considerations of constitutional variability. Although handedness is commonly treated in terms of a simple dichotomy, there is good reason to believe that lateral dominance is in reality a graded characteristic and that constitutional differences exist in both the strength and the consistency of lateral preference. In our present cases, it might be supposed that the degree of inborn sinistrality was relatively weak, leading to a lack of clear unilateral hand preference and to certain inconsistencies in preference as between hand, foot and eye. As a result, one might expect hand usage to have evolved in almost random fashion, being determined to a far greater degree than is usual by incidental factors, either personal or environmental. For instance, one subject (No. 6) reported that she tended to prefer her right hand for activities involving muscular effort (e.g. throwing) and her left for activities requiring close guidance (e.g. sewing or billiards). Although most individuals tend to use the preferred hand irrespective of considerations of relative strength, it is possible that

a dissociation of this kind is not uncommon among those who are poorly lateralized. Another subject (No. 1) remarked that she thought she had been more strongly left-handed as a child but had gradually become less so as she grew up. She could not recall any conscious wish or intention to shift her hand preference and quite certainly had not been subjected to pressure by others to do so. This might suggest delayed maturation of lateral dominance in a virtually ambilateral (and somewhat "ambilaevous") individual. One may also point out that some slight admixture of left-hand preferences is by no means uncommon among right-handed individuals with left-handedness in the family history. This can hardly be explained in terms of social training and is presumably due to an obscure genetical influence.

In the light of these considerations, we would suggest that atypical handedness patterns of the type described are only indirectly related to factors of training and social convention. We believe that a basic sinistral tendency, linked with some constitutional factor, was present in all ten subjects, but that this tendency was insufficiently marked to constrain strong and consistent unilateral preferences. In consequence, hand usage showed far greater variability than is usual and inconsistencies of preference were strongly marked. This finding may perhaps be related to recent neurological evidence suggesting that unilateral cerebral dominance is often far from clear-cut in individuals who are partly left-handed or who display mixed patterns of hand preference (Humphrey and Zangwill, 1952; Ettlinger, Jackson and Zangwill, 1956). At all event, it would appear to warrant reconsideration of the older concept of the "ambidexter" as a basic biological variant.

References

Blau, A. (1946). "The master hand." *Amer. Orthopsychiat. Ass. Res. Mon.* No. 5.

Brain, W. Russell (1945). Speech and handedness. *Lancet*, **2**, 837–42.

Burt, C. (1950). *The Backward Child.* London.

Ettlinger, G., Jackson, C. V., and Zangwill, O. L. (1956). Cerebral dominance in sinistrals. *Brain*, **79**, 569–88.

Hildreth, G. (1949). The development and training of hand dominance. *J. genet. Psychol.*, **75**, 197.

Humphrey, M. E. (1951). Consistency of hand usage. *Brit. J. educ. Psychol.*, **21**, 214–25.

Humphrey, M. E., and Zangwill, O. L. (1952). Dysphasia in left-handed patients with unilateral brain lesions. *J. Neurol. Neurosurg. Psychiat.*, **15**, 184–93.

Scheidemann, N. V., and Colyer, H. (1931). A study of reversing the handedness of some left-handed writers. *J. educ. Psychol.*, **27**, 191.

Wile, I. S. (1934). *Handedness: Right and Left.* Boston.

THE EFFECT OF TEMPORAL SEPARATION AND PRIOR HABITUATION UPON PRECONDITIONING*

BY

EDGAR HOWARTH

From the Department of Psychology, University of Melbourne

The effect of temporal separation and prior habituation was tested by measuring response transfer in a shuttle box following different initial treatments. Thirty hooded rats were split into three main groups and six sub-groups. One main group was preconditioned with concurrent light and sound stimuli, another with temporally spaced stimuli, and a control group had no preconditioning. Half of each main group was permitted to explore the apparatus before preconditioning on the supposition that this might reduce emotionality and hence reduce startle or attention responses.

It was found that concurrent stimulation produced significantly greater response transfer than in controls, whereas preconditioning with spaced stimuli did not. Prior habituation produced no significant effect on the amount of response transfer.

Introduction

This experiment was undertaken shortly after the appearance of Reid's (1952) criticism of Brogden (1939). In view of the theoretical importance attached to preconditioning by Birch and Bitterman (1949) it seemed to be of some urgency that the status of the phenomenon should be further examined. Was preconditioning a genuine conditioning phenomenon or did it arise from an inadequacy in Brogden's experimental design? If preconditioning were genuine it could be expected to bear some relationship to conventional conditioning parameters. As studies of classical conditioning (Wolfle, 1932, Razran, 1957) had shown that temporal separation of the CS and the UCS was an important variable, it was decided to examine the effect of temporal separation of the preconditioning stimuli. It should be noted that at the time of commencement of this experiment the results of Silver and Meyer (1954), which show the importance of temporal order of stimulation, were not available.

When Brogden (1939) introduced the term "sensory preconditioning" he did not cite the experiment of Narbutovitch and Podkopayev (cited by Konorski, 1948) which led to the supposition that paired stimuli could become associated on the basis of an orientation response. In a recent paper on the neurophysiology of conditioning Gastaut *et al.* (1957) have strengthened this view. In designing the present experiment the effect of prior habituation to the apparatus was tested in order to see whether this might reduce emotionality during preconditioning and, hence, startle or attention responses made to the preconditioning stimuli.

If either of the preconditioning stimuli acted as a CS for an orientation response to the other stimulus sufficient temporal separation could be predicted to annul the effect of preconditioning. As previous studies had shown that temporal separation of CS and UCS greater than 4 sec. made conditioning negligible, it was decided to separate the onset of the stimuli by an amount greater than this. Following Brogden's general procedure, after preconditioning its effect was assessed by conditioning

* Based on a paper presented to the 13th Annual Conference of the British Psychological Society (Australian Branch), Sydney, August, 1957.

the animal with one of the preconditioning stimuli (light) then the remaining stimulus (sound) was substituted in test trials. As the shuttle-box procedure seemed to offer a convenient method of conditioning the animals it was planned that, after preconditioning, an avoidance response would be conditioned using light as CS, and then sound would be substituted as CS in further training. Experimental groups preconditioned with or without temporal separation of stimuli during preconditioning could then be compared with each other, and with a control group lacking preconditioning, on the amount of transfer shown by comparing the percentage saving in number of responses required to attain the learning criterion using sound as CS with the number required using light as CS. It was predicted that when the stimuli were temporally concurrent during preconditioning most saving would occur, that the least saving would occur in the control group, and that the temporally spaced group would not differ significantly from the control group. With regard to prior habituation it was predicted that this would reduce the effect of preconditioning in the experimental groups and, in particular, in the concurrent group.

Method

Apparatus

A shuttle-box was used, of interior dimensions 24 in. long, 12 in. wide, 11 in. high, the floor of which consisted of two separately wired electrifiable grids, each 12 in. square. These were separated by a central hurdle 2 in. high, running across the width of the apparatus. A brief shock (0·25 Ma.) was applied alternately to each grid during training procedures. The top of the hurdle was constantly electrified during conditioning, but not during habituation and preconditioning. The light stimulus consisted of a 6 V. lamp centrally mounted, facing downward into the box. The sound stimulus consisted of a pure tone 1000 c.s. of 40 db. intensity (reference level 0·0002 dynes per sq. cm.) emanating from a loudspeaker mounted in the centre of one wall of the box facing into the box.

Timing relationships

During preconditioning, for the temporally concurrent group, light and sound stimuli of 2 sec. duration were regularly paired at 15 sec. intervals; for the temporally spaced group light and sound stimuli were presented individually, the onset of stimuli being spaced by 7·5 sec. During conditioning a 2 sec. CS was delivered with its cessation co-incident with the onset of the UCS (shock.)

Subjects. Thirty experimentally naïve hooded rats, 2 to 4 months old were used as subjects.

Procedure. After initial handling for 2 weeks, half of the subjects randomly selected, were allowed to explore the shuttle box for three 1-hr. periods on three successive days, the last period being on the day preceding preconditioning. Subjects were further divided into three main preconditioning treatments:

The *temporally concurrent* group was given 500 presentations of paired light and sound stimuli in one 125 min. period. The *temporally spaced* group was given 500 presentations of light and sound which, although delivered regularly, were at no time contiguous or concurrent. The control group was simply enclosed in the apparatus for 125 min. without experiencing the preconditioning stimuli. All animals were preconditioned individually.

Immediately following the above treatment the animal was trained to avoid shock, using the light as CS. The criterion was a total of 10 successful avoidance trials, i.e. jumping over the hurdle within 2 sec. following onset of the CS. In the final phase sound was immediately substituted for light as CS and the amount of transfer was estimated by training the animal to the same learning criterion as in the previous phase.

The procedure may be summarized by stating that the animals were first habituated as necessary, then preconditioned as necessary, then conditioned using light and, finally, conditioned using sound.

Results

The percentage saving scores obtained by this method are shown in Table I. A positive percentage saving resulted if the subject attained the learning criterion with sound as CS in less trials than with light as CS. Analysis of variance applied to this data revealed that the overall difference between groups was significant ($F = 4{\cdot}97$ with 2 and 24 d.f.) whereas the effect of prior habituation was not significant ($F<1$ with 1 and 24 d.f.). Application of the Fisher-Yates exact test (Finney, 1948) to the data revealed that whereas the concurrent group differed significantly from the controls ($p<0{\cdot}05$), the spaced group did not so differ ($p>0{\cdot}05$), furthermore, the two experimental groups differed significantly ($p<0{\cdot}05$).

TABLE I

Percentage Saving using Shuttle-Box Method*

	Concurrent group	*Spaced group*	*Control group*
Experienced Subjects†	+6·2	−9·5	−30·1
Naïve Subjects	+29·4	−6·6	−54·0

* Light was used first as CS and finally sound as CS. If less trials were needed to attain the learning criterion using sound the saving percentage was positive.

† $n = 5$ in each cell.

Discussion

These results appear to indicate that separating the onset of two preconditionng stimuli (light and sound) significantly reduces the effectiveness of preconditioning. At the same time, if preconditioning were due, as Reid (1952) suggested, merely to the fact that experimental subjects had more experience of the test stimulus than controls prior to the test phase, it is difficult to see why subjects preconditioned with concurrent stimuli performed better than subjects preconditioned with temporally spaced stimuli, as both of these groups had equal experience of the sound stimulus prior to the test phase. One is therefore disposed to believe in the genuine nature of the phenomenon.

As to the basis of the phenomenon, the absence of an effect due to prior exploration appears to show that an emotional (startle) response does not necessarily mediate preconditioning but the amount of prior exploration may have been insufficient or the interval between the last period of exploration and preconditioning may have been too long.

In summary: This experiment validates the status of preconditioning in two ways: (*a*) it indicates that Reid's criticism is unwarranted, which points to the possibility that his negative experimental finding may have been due to an insensitive procedure; (*b*) it shows that preconditioning is reduced by a temporal spacing of CS and UCS, and is therefore sensitive to a procedure which has been shown to effect classical conditioning.

References

Birch, H. G., and Bitterman, M. E. (1949). Reinforcement and learning: the process of sensory integration. *Psychol. Rev.*, **56**, 292–308.

Brogden, W. J. (1939). Sensory preconditioning. *J. exp. Psychol.*, **25**, 323–32.

Finney, D. J. (1948). The Fisher-Yates test of significance in 2 × 2 contingency tables. *Biometrika*, **35**, 146–56.

Gastaut, H., *et al.* (1957). Étude topographique des réactions électroencéphalographiques conditionnées chez l'homme. *EEG Clin. Neurophysiol.*, **9**, 1–34.

Konorski, J. (1948). *Conditioned Reflexes and Neuron Organization.* London.

Razran, G. (1957). The dominance-contiguity theory of the acquisition of classical conditioning. *Psychol. Bull.*, **54**, 1–46.

Reid, R. L. (1952). A test of sensory preconditioning in pigeons. *Quart. J. exp. Psychol.*, **4**, 49–56.

Silver, C. A., and Meyer, D. R. (1954). Temporal factors in sensory preconditioning. *J. comp. physiol. Psychol.*, **47**, 57–59.

Wolfle, Helen M. (1932). Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, **7**, 80–103.

APPARATUS

A NEW COLOUR MIXER AND EPISCOTISTER

BY

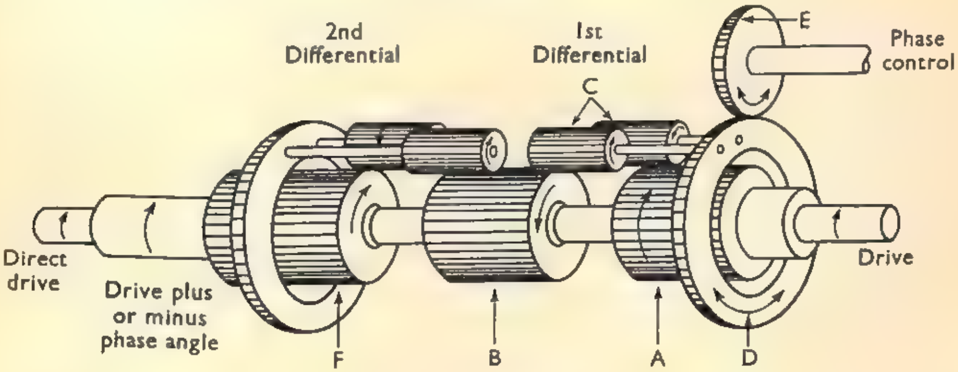
A. PERRY and I. P. HOWARD

From the Department of Psychology, University of Durham

The paper describes the construction of a differential system which can be used as a colour mixer or flicker apparatus. The use of sectored discs has been largely replaced by filters, wedges and polaroids in much visual work. However, our apparatus has two main applications, as a convenient, inexpensive colour mixer for student work and as a research tool in the study of flicker fusion and kindred phenomena.

Most psychology laboratories would find use for a colour mixer which gives a continuous change in the proportion of one colour disc to another while the colour wheel is running. Such a device with sectored discs is also useful as a convenient and precise method of controlling light intensity or as a visual flicker apparatus with a variable on–off ratio. Commercially manufactured differential colour mixers are expensive and bulky. A less precise, though cheap, design was described by Young (1923). The present design combines cheapness and compactness with precision and relative ease of manufacture.

FIGURE 1



The system is based on the differential gear. In such a system a gear on an input shaft (A)* drives a gear (B) in reverse through two intermediate gears (C) The axles of the intermediate gears are attached to a housing with gear teeth on its periphery (D). A rotation of the housing through a given angle imparts a phase shift of twice this angle between the input and output gears. The control to D is through a 2 : 1 reduction gear (E) which enables the correct phase angle to be indicated on a 360° dial. A spring friction-washer on the shaft of gear E prevents the phase from shifting of its own accord.

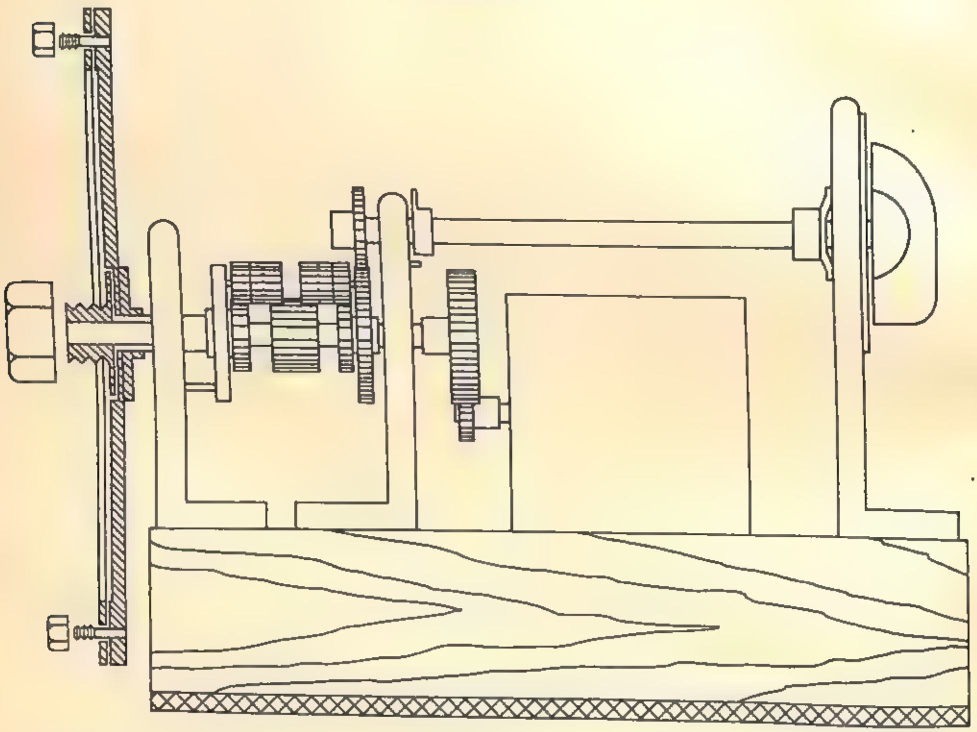
* The letters in the text refer to those on Figure 1.

The two shafts from a single differential rotate in opposite directions. The output gear from this first differential may be reversed by passing it into a second differential in line with the first.

The original drive shaft is brought through the hollow centre of the entire system to give two concentric shafts rotating in the same direction upon which may be amounted the coloured discs. The housing of the second differential (F) is held stationary.†

A gear coupling between motor and the input shaft is desirable to allow for any small misalignment and to give a step-down or step-up gear ratio if required. Small ball-race bearings make for smooth running.

FIGURE 2



The disc mounting is shown in section (Fig. 2) and allows about 350° phase shift; a small angle must be left so that the papers remain interleaved. The first colour disc is attached to the centre shaft emerging from the differentials by a half in. nut on a boss recessed into a $\frac{1}{8}$ in. thick plastic disc 6 in. in diameter. The second paper disc is attached to the plastic disc by a large plastic ring and thumb screw. The plastic disc is also recessed $\frac{1}{32}$ in. on its front face to the diameter of the first colour disc to allow the two paper discs to move freely relative to one another. The hole in the second paper disc and the outside edge of the first paper disc must be cut away slightly so that the discs avoid each other's points of attachment.

Sectored discs are easily attached for the mechanism to be used as a flicker apparatus rather than a colour wheel. A thick wooden base with sponge rubber on its underside gives stability and relative freedom from noise.

† The differentials are obtainable from H. Franks, 58–60 New Oxford Street, W.C.1, Catalogue item No. 97, for a few shillings.

We have coupled two such systems together by a third single differential. With this we are able to stimulate each eye separately at the same frequency and vary the on–off ratio of flicker independently to each. At the same time the phase of the flicker to one eye may be varied relative to that of the other. This is very much cheaper and less bulky than the equivalent electronic circuitry. A very accurate frequency is maintained by driving the system with an oscillator controlled Desyn motor, ½ in. in diameter. Using small differentials from old coin-slot meters, this flicker apparatus occupies the volume of a 4 in. cube.

Reference

Young, P. T. (1923). A differential colour mixer. *J. exp. Psychol.*, **6**, 323–43.

BOOK REVIEWS

Verbal Behavior. By B. F. Skinner. London: Methuen. 1959. Pp. x + 478. 42s.

Skinner tells us that this book had its origins in a challenge put to him by Whitehead in 1934. They became engaged in a prandial discussion about behaviourism. Whitehead agreed that "science might be successful in accounting for human behavior provided one made an exception of *verbal* behavior. Here, he insisted, something else must be at work. He brought the discussion to a close with a friendly challenge: 'Let me see you,' he said, 'account for my behavior as I sit here saying "No black scorpion is falling upon this table".' The next morning I drew up the outline of the present study."

How does Skinner meet the challenge? By giving us what he calls "a functional or causal analysis" of verbal behaviour. By the latter term he means, at the outset and provisionally, "behavior reinforced through the mediation of other persons." "The basic facts to be analysed," he tells us, "are well known to every educated person." By "functional analysis" he means in effect a description of these facts in terms of his own behaviour theory. "The emphasis [of the analysis]," he writes, "is upon an orderly arrangement of well-known facts, in accordance with a formulation of behavior derived from an experimental analysis of a more rigorous sort. The present extension to verbal behavior is an exercise in interpretation rather than a quantitative extrapolation of rigorous experimental results." He argues that the traditional formulations and concepts are of little help in attempting to arrive at a functional analysis. Thus he finds no use for concepts such as word, meaning, proposition, sign, symbol, intention of the speaker, and the like. Consequently, he has to devise new concepts of his own. The upshot is a long book containing a number of unfamiliar terms, in which the field of verbal behaviour is systematically described in Skinnerian language.

Let us pick out a few of the details in this grand exercise. Skinner begins by noting the chief features of verbal behaviour as a dependent variable. The unit that he selects is "an operant," but this cannot be identified by its form alone. For *x* to be an instance of a verbal operant, it must also be a function of some certain variable. The first type of verbal operant he distinguishes is "the Mand." This is one "in which the response is reinforced by a characteristic consequence and is therefore under the functional control of relevant conditions of deprivation or aversive stimulation." For example: "Listen!" "More soup!", "Pass the salt!" He then distinguishes three types of verbal operant that are under the control of verbal stimuli. These are: "Echoic Behaviour"; "Textual Behaviour" (for example, when a person is reading); and "Intraverbal Behaviour" (for example, when I say "Four" to the stimulus "Two plus two"). Next, he distinguishes "The Tact." This is "a verbal operant in which a response of given form is evoked (or at least strengthened) by a particular object or event or property of an object or event. We account for the strength by showing that in the presence of the object or event a response of that form is characteristically reinforced in a given verbal community." Here Skinner is obviously concerned with verbal behaviour in which we "talk about" the world. He goes on to describe "the Extended Tact" (in metaphor, for instance) and to deal with "Abstraction" and "the Problem of Reference." He discusses the role of the other important type of controlling stimulus that is also non-verbal, namely the audience; and he has a separate discussion of "the verbal operant as a unit of analysis," ending up this part of the book with a refinement of his definition of verbal behaviour. He continues in the next part by discussing some of the ways in which the controlling variables interact.

So far, however, he points out, he has assumed that the speaker does not know what he is saying, or wants to say, or how to say it. But "order, design and 'deliberate' composition are observable features of verbal behavior." Now these become possible, Skinner argues, because the speaker acquires a system of behaviour that controls his own behaviour; and Skinner uses this notion to introduce another new term, "the "Autoclitic." This term is "intended to suggest behavior which is based upon or depends upon other verbal behavior." He distinguishes various types of autoclitics. For example: (*a*) descriptive, where one "tacts" one's own verbal behaviour (for instance, "I believe . . .", "I surmise . . .", which describe the state of strength of a response; "I am happy to say . . ." indicates the emotional condition of the speaker; "I demand . . .

tells the listener what kind of verbal operant it accompanies, namely a mand); (*b*) qualifying autoclitics (for instance, negation—"No," and assertion); (*c*) quantifying autoclitics (for instance, "a," "the"); and (*d*) those that are the traditional concern of grammar and syntax (such as inflections, the order in which responses appear, and predication). Skinner is at pains to try to show how the autoclitic activity of the speaker enables him to manipulate and control his own verbal behaviour, especially in composition. But, in addition to this, the speaker also examines his verbal responses for their effect on himself or listener, and they are then either rejected or released. Skinner calls this the "process of 'editing'." The last part of the book is taken up with a discussion of this process, and it ends with brief descriptions of logical and scientific verbal behaviour, and of thinking.

How, then, does Skinner answer Whitehead's challenge? The remark "No black scorpion is falling on this table" was meant to be a poser just because there was no black scorpion falling on the table—just because, that is, it was obviously *not* controlled by a present stimulus. But in a determined system there must have been some stimulus at work, even if it only weakly determined the form of the response. Now, just as a physicist at the present time may suggest various explanations of the fact that, say, there was a sudden drop in the temperature of the room at that dinner in 1934 (in order to show that the occurrence can be explained lawfully), so it is not out of place for him, Skinner, to make a guess today about Whitehead's remark. "I suggest, then," says Skinner, "that 'black scorpion' was a metaphorical response to the topic under discussion. The black scorpion was behaviorism." He continues a little further on: "It is possible, then, that as I described my position—doubtless in the most shocking terms I could command—he [Whitehead] was telling himself that the part which he had played in encouraging me as a young scholar was not entirely misguided, that I was probably not typical of all young men in psychology and the social sciences, that there *must* be a brighter side—in other words, that on this pleasant and stimulating table no black scorpion had fallen." Skinner adds that this explanation "is, of course, only a most improbable guess."

What are we to make of all this?

(1) To the outsider it is not immediately obvious how "the functional analysis" contained in this book will be of much help to the experimental psychologist—no matter whether the latter works inside or outside Skinnerology. If Skinner wanted the experimentalist to get something out of this book, he could have helped him by showing where the edifice can be brought down and tied to the earth of the laboratory. As it is, he scarcely mentions any experimental work, past or present.

(2) Is it logically necessary to spin this colossal story in order to answer Whitehead's challenge? No, it is not. Skinner himself makes this quite clear by his own answer: this is logically independent of his own functional analysis.

(3) Why, then, did he spin this story? He seems to have been trapped into supposing that Whitehead's question was one *inside* psychology, and not one *about* it. I think this supposition is a mistake. But given that he took the question as being one inside psychology, he naturally proceeded to answer it as a psychologist; and *in America* and *at that time* he could only be expected to answer it in one way, namely, by producing a large scale language which enables us to talk in a systematic way about the whole field of verbal behaviour. The book is to be understood, therefore, as a product of American culture—and perhaps, also, as a period piece.

(5) What is its value? Quite apart from any stimulus that the experimentalists may extract from it, it is of value in at least two different directions. It is an important exercise in theoretical psychology. Skinner has shown us what is involved in extending to verbal behaviour the apparatus of his behaviour theory. It is the first time, to my knowledge, that any leading theorist has attempted to talk in detail about linguistic functioning, and we should be grateful to him for his efforts. In making this attempt, he has brought closer together the two disciplines of psychology and logic. Here the book may be of help to many logicians and philosophers. It is time that the bogey of psychologism was slain, and a move made towards the construction of what Skinner calls "an empirical logic." He has probably given us some of the weapons necessary for this job; and students in this field will certainly derive a great deal from the insights and the analysis that this book contains.

B. A. FARRELL.

The Bristol Social Adjustment Guides. By D. H. Stott and E. G. Sykes. London. University of London Press. 1958. Guides: Specimen Set 3s. 10d. Manual: *The Social Adjustment of Children.* Pp. 46. 25s.

The purpose of these "Guides" is the recognition of type and degree of emotional instability among children from 5 to 15 years of age. They cater for three different situations: the child in day-school, in residential care and in the family, the proformas being completed by school-masters and -mistresses, staff-members of residential centres and social workers, respectively. In each case, the Guide consists of four printed pages, giving brief, clear descriptions of different types of behaviour, classified under major headings.

For instance, the school guide includes items on Attitude to Teacher, to School Work, to other Children, Games and Play, Personal Ways and Physique. The guide for children in residential care covers some of the same topics but includes further questions on attitude to relatives, sleep habits, etc. The third proforma, the Family Guide, includes a life history chart and has sections on parent-child relationship, behaviour in home, health and development. The assessor is instructed to check one or more items that apply to the relevant child but to omit where no item is appropriate. He is asked to add supplementary comments where necessary. The items throughout are scrambled, in order to minimise halo effect; and most paragraphs include at least one "normal" response. On scoring the guide, a pattern emerges which gives some idea of the direction, complexity and extent of the child's divergence from normality—if divergence there be. These results are expressed in terms of, for example, Anxiety (X), towards adults (XA) or other children (XC); Depression (D); Withdrawal (W) or, less severe, Unforthcomingness (U); Hostility (H) or Hostility-Anxiety (HX); Knavery (K), which appears to be psychopathic anti-social behaviour; and so on.

Reliability is remarkably high for this field, correlations between pairs of teachers, for instance, being 0·78 for "unsettled scores" and 0·76 for "maladjusted scores." Validation has been achieved mainly inductively by comparing individual test-items and also "test-syndromes" with various predetermined groups, and repeatedly checking and improving the relationship. The authors do not, however, lose sight of the individual child in their appreciation of objectivity and of systematic research. They write, for example, that "the temptation to state the diagnosis in terms of a general quotient has been resisted. Any quotient or 'score' for maladjustment would be artificial because there are different ways of being maladjusted which are hardly comparable to one another." They withstand also the attractions of a typological approach, of statistics as an end in itself and of personality and trait psychology.

Inevitably, a number of criticisms suggest themselves. There is, for instance, too much philosophizing and self-interested history in the Section called "Methodology and Compilation"; there is, also, an infuriating tendency to assume familiarity on the part of the reader with the particular jargon chosen; and certain sections of the Manual suggest that, despite six years' work, publication in this form may be slightly premature. In general, however, the Guides are impressive. Stott and Sykes have tackled a perplexing problem without underestimating its difficulties and without oversimplifying method or interpretation. It is one of the most hopeful attempts yet made in this important field.

A. W. Heim.

Figural After-effects. By Peter McEwen. (*British Journal of Psychology: Monograph Supplement,* No. 31.) London. Cambridge University Press. 1958. Pp. viii + 106. 22s. 6d.

A review of an experimental field, if it is adequate, may reasonably be expected to reveal deficiencies in experiment, to put controversial theoretical issues into sharper relief and so, directly or indirectly, lead to ideas for further experiment or to new theoretical questions. This monograph on figural after-effects meets these criteria in a satisfactory way.

Pride of place is given, rightly, to an account of J. J. Gibson's work as having been the first systematic description and investigation of visual phenomena of the kind which subsequently came to be called figural after-effects. But experimental work on after-effects, as the author states, has been dominated by the monumental series of effects described by Köhler and Wallach and, especially, by the satiation theory which they erected to account for them and to relate them to other visual phenomena. One third of the monograph is therefore devoted to a critical exposition of the work of Köhler and

his collaborators in relation to later work. Included are their investigations of visual after-effects in the third dimension, of "kinaesthetic" effects analagous to displacements occurring in visual after-effects, and the attempt to give an account in terms of satiation of the destruction of illusions when repeatedly inspected. The examination is thorough and attention is directed to critical instances in which either observations reported appear at variance with expectation from the theory, or *ad hoc* extensions are made to the theory to account for them. The alternative "statistical" theory of Osgood and Heyer, to which another chapter is devoted, is treated similarly. What do perhaps emerge most clearly from this systematic exposition are the obstacles to rigorous deduction from either theory. Quantitative deductions, at least of relative magnitudes, are required. But the premises of both theories lack precision, argument from them has to be by circumlocution and an uncertainty of conclusion is the inevitable result. In the case of the statistical theory, the questions of the shape and scale of the distributions hypothesized are important if not vital. Neither question was considered in any detail by Osgood and Heyer. The basis of their theory in Marshall and Talbot's theory of visual acuity might with advantage therefore have been subjected to a more penetrating scrutiny. Deutsch's valuable, if negative, contribution on these points appeared when much of the monograph was written and the references to it are rather brief. Other data on acuity remain unconsidered, however, and their examination in relation to after-effects may yet prove fruitful.

A fourth chapter is devoted to other experimental studies. Included are reviews of work on kinaesthetic and auditory after-effects, as well as on visual effects in relation to apparent size, apparent movement and reversible figures. It concludes with the author's cautious appraisal of the present position. Generally, the monograph succeeds in giving a perspective to its topic. It does so even though some salient questions are touched upon, but not pursued. It provokes, rather than answers, such questions as the sense in which figural after-effects are distinguishable from other perceptual phenomena, how far the description "figural" is applicable to all of the after-effects so described, particularly as between sense modalities, and what might be their relation to other kinds of perceptual distortions and adjustments. That these, and other questions calling more immediately for further experiment, should be emergent, is nevertheless a measure of the value of this review.

ARTHUR SUMMERFIELD.

La Magie du Dessin. By Jean Vinchon. Bibliothèque neuro-psychiatrique de langue française. Bruges. Desclée de Brouwer. 1959. Pp. 182. 150 FRB.

Die Sprache der Zeichung. By Carlos J. Biedma and Pedro G. D'Alfonso. Bern. Huber. 1959. Pp. 110. Fr/DM 24.80.

Dr. Vinchon's book is a study of "doodling" and its supposed psychological significance. Although not experimental, it contains much interesting material, normal and abnormal, and a fairly level-headed commentary. Some discussion of the place of art in psychotherapy is attempted.

Dr. Biedma and Dr. D'Alfonso give an account of their new version of the Wartegg drawing test, which is supposed to have value in diagnostic personality study. They present a detailed analysis of the various features of style and expression elicited by the test, and claim some validity for their interpretations. The method may be of some interest to clinical psychologists of projective inclination.

O. L. ZANGWILL.

CORRIGENDUM

The ratio on p. 198 (line 3) and on p. 201 (line 8) of the article by A. T. Welford in the November, 1959 issue of this *Journal* should read 1/1.59 instead of 1.59/2.

PUBLICATIONS RECEIVED

NEW PERIODICALS

The following publications have been received:—

Activitas Nervosa Superior. Vol. I, Nos. 3 and 4, 1959. Chief Editor: M. Horváth. (Published by the Society for the Study of the Higher Nervous Activity, Section of the Czechoslovak Medical Society, J. E. Purkyně, Prague). Obtainable from Artia, Prague 2, Ve Smečkách 30. English Summaries.

Grundlagenstudien aus Kybernetik und Geisteswissenschaft. Vol. 1, No. 1, January, 1960. Edited by M. Bense, F. von Cube, G. Eichhorn, H. Frank, A. A. Moles and E. Walther. (Technische Hochschule, Huberstrasse 16, Stuttgart.)

Izobrazba Rukovodilaca. No. 2, June, 1958. (Zagreb, TRG Žrtava Fašizma 4.) Industrial Psychology and Management Studies. English summaries.

Journal of Primatology: Primates. Vol. 1, No. 2, 1958. Edited by K. Andô, D. Miyadi, Y. Tajima and K. Imanishi. (Published by the Japan Monkey Centre, Kurisu Inuyama City, Aichi, Japan.) Japanese and English text.

Rehabilitation. (Section XIX of *Excerpta Medica.*) Vol. 1, No. 1, July, 1958. Edited by M. Fishbein, M. W. Woerdeman, and H. Ogilvie. (111, Kalverstraat, Amsterdam.)

OTHER PERIODICALS RECEIVED

Periodicals mentioned in the list published in this *Journal*, 1958, **10**, 176, are not included.

Archives Italiennes de Biologie. Vol. 97, No. 1, 1959. Edited by G. Moruzzi. (Pisa Universita Degli Studi, Physiological Institute of the University of Pisa, Via S. Zeno No. 11–13, Pisa.)

Japanese Journal of Psychology. Vol. 30, No. 4, 1959. Edited by M. Sagara and Y. Togawa. (Published bimonthly by the Japanese Psychological Association, c/o Faculty of Letters, University of Tokyo, Tokyo.) Japanese text with some English summaries.

Japanese Psychological Research. No. 8, 1959. Edited by T. Obonai and M. Sagara. (Published by the Japanese Psychological Association, c/o Faculty of Letters, University of Tokyo, Tokyo.) English text.

Psychologia Wychowawcza Kwartnalnik (Educational Psychology). Edited by M. Zabrowska. (Warsaw, ul. Spasowsklego Nr 6/8.)

Psychologie Française. Vol. 4, No. 3, 1959. Revue trimestrielle de la Société Française de Psychologie. Directed by P. Fraisse. (Société Française de Psychologie, 46 Rue St. Jacques, Paris 5e.)

Soviet Information Bulletins: Psychology. Vol. 6, Nos. 1 and 2, 1959. Price 3s. (Society for Cultural Relations with the U.S.S.R., 14 Kensington Square, London, W.1.)

OTHER PUBLICATIONS RECEIVED

Charles Darwin. By Sir Gavin de Beer. Henriette Hertz Trust Lecture of the British Academy, 1958. Oxford University Press, 1958. Price 3s. 6d.

Manual of the Maudsley Personality Inventory. By H. J. Eysenck. University of London Press, 1959. Price 2s. 6d. net.

Tobacco Manufacturers' Standing Committee: Research Papers No. 2. The Reliability of Statements about Smoking Habits. By G. F. Todd and J. T. Laws. (6–10, Bruton Street, London, W.1.)

PERGAMON INSTITUTE

The Pergamon Institute (English address: Headington Hill Hall, Oxford) announce that English editions of the following journals are being prepared: *Pavlov Journal of the Higher Nervous Activity*, *Problems of Psychology* (selected papers) and *Problems of Cybernetics.* Specimen copies and subscription rates on application. A series of monographs in translation, including some psychological titles, is also in preparation.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XII 1960 Part 3

ON THE FAILURE TO ELIMINATE HYPOTHESES IN A CONCEPTUAL TASK

BY

P. C. WASON

From the Medical Research Council Industrial Psychology Research Group, University College, London

This investigation examines the extent to which intelligent young adults seek (i) confirming evidence alone (enumerative induction) or (ii) confirming and disconfirming evidence (eliminative induction), in order to draw conclusions in a simple conceptual task. The experiment is designed so that use of confirming evidence alone will almost certainly lead to erroneous conclusions because (i) the correct concept is entailed by many more obvious ones, and (ii) the universe of possible instances (numbers) is infinite.

Six out of 29 subjects reached the correct conclusion without previous incorrect ones, 13 reached one incorrect conclusion, nine reached two or more incorrect conclusions, and one reached no conclusion. The results showed that those subjects, who reached two or more incorrect conclusions, were unable, or unwilling to test their hypotheses. The implications are discussed in relation to scientific thinking.

Introduction

Inferences from confirming evidence (Bacon's "induction by simple enumeration") can obviously lead to wrong conclusions because different hypotheses may be compatible with the same data. In their crudest form such inferences are apparent in the selection of facts to justify prejudices. In research merely confirming evidence is clearly of limited value. For example, suppose that it is suggested that a deficit (x) of a particular substance in the blood is uniquely related to a distinctive symptom (y). (In logical terms, x is a "necessary-and-sufficient" condition for y.) And suppose that this hypothesis had been supported by confirming evidence alone, i.e. whenever the deficit had been induced, the symptom had appeared (ignoring statistical issues for the sake of simplification). It might then be assumed that the hypothesis was tenable. But the evidence only allows the inference that x is a sufficient condition for y, that the deficit always leads to the symptom. To establish the postulated relation, there must also be no disconfirming evidence—no case of the symptom without the deficit. For if such a case were obtained, the deficit would not be a necessary condition for the symptom, and the symptom would not be a sufficient condition for the deficit. The hypothesis that either was a necessary-and-sufficient condition for the other could be eliminated. The symptom would not be a reliable sign. Its absence would rule out the possibility of the deficit, but its presence would be ambiguous.

In general, scientific inferences are based on the principle of eliminating hypotheses, while provisionally accepting only those which remain. Methodologically, such eliminative induction implies adequate controls so that both positive and negative experimental results give information about the possible determinants of a phenomenon.

This investigation seeks to determine the extent to which intelligent young adults make rational inferences about abstract material which does not obviously suggest a conventional scientific approach. A concept attainment task would seem to be most suitable for this purpose, and one such typical task will now be considered in detail.

Bruner, Goodnow and Austin (1956) report some ingenious experiments of this kind. The material of one experiment consisted of an array of 81 instances made up from all the combinations of four attributes, each exhibiting one of three values, i.e. shape (cross, square, circle), colour (green, red, black), number of figures (one, two, three) and number of borders (one, two, three). The subject's task was to attain a concept, defined in terms of the values of these attributes, by choosing successive instances in order to find out whether they exemplified the concept. The authors describe a number of possible strategies which can be used. One of these, "successive scanning," is equivalent to induction by simple enumeration in that it involves testing successive hypotheses by using information derived from positive instances alone. For example, if "green squares" is the subject's hypothesis, instances exhibiting these features are selected. If these turn out to be positive, "green squares" is considered to be necessary-and-sufficient for the concept. But if any of these instances are negative, a different hypothesis is similarly tried out, and accepted so long as it leads to positive instances.

However, Bruner *et al.*, are not concerned with the fact that successive scanning can lead to merely sufficient, as opposed to necessary-and-sufficient, concepts. For example, if the correct concept is "red figures" (of which there are 27 positive instances), it is possible to attain the incorrect concept, "red circles" (of which there are 9 positive instances), by consistent use of confirming evidence, since all positive instances of this latter concept are also positive instances of the correct one. The incorrect concept, in this case, entails the correct one but is not entailed by it, i.e. "red circles" is a sufficient, but not a necessary-and-sufficient, condition for "red figures." Thus it is logically possible to arrive at incorrect concepts by successive scanning because only confirming evidence is utilized. Such a result, however, may not occur often because other instances, which might act as reminders of alternative possibilities, are displayed in front of the subject. In the above example, the subject's attention might be directed to the possibility of "circles" being necessary, owing to the presence of instances exhibiting red squares and crosses.

The present investigation is designed to compel the subject to encounter plausible sufficient conditions of the concept which must ultimately be eliminated in order to attain the necessary-and-sufficient conditions. The logical mechanism underlying efficient performance in the task rests upon what Von Wright (1951) terms "the fundamental, though trivial, fact that no confirming instance of a law is a verifying instance, but that any disconfirming instance is a falsifying instance."

In the experiment the concept to be attained is "*three numbers in increasing order of magnitude.*"

Subjects were told that the three numbers 2, 4, 6, conformed to a simple relational rule and that their task was to discover it by making up successive sets of three numbers, using information given after each set to the effect that the numbers conformed, or did not conform, to the rule.

It will be seen that this task, as a whole, differs from previous studies of concept attainment. In the first place, the attributes are rules referring to relations between numbers, e.g. "consecutive even numbers." In this respect, the possible concepts are more like those found in the Goldstein-Scheerer (1941) test of abstractive ability, than those which can be attained in Bruner's experiment, in which attributes cannot

be combined into classes designated by names but have to be enumerated, e.g. "three red circles."

Secondly, the possible instances (triads of numbers) are, in principle, infinite. In previous studies a finite universe of instances has been used. Such a universe, however, places great constraint on the scope of inductive thinking because the number of instances which exemplify any sufficient concept will always be less than the number which exemplify any necessary-and-sufficient one. Hence the number of instances of a given kind which can be tested is limited. But in the present task an endless series of instances, exemplifying a sufficient rule can be generated without forcing the subject to encounter an instance which would not exemplify it.

Thirdly, the instances are not presented as stimuli, but have to be generated by the subject. In this way he is completely free to decide on the kind and amount of evidence which he considers adequate. If, on the other hand, all the possible instances are displayed simultaneously, the subject will know that all available evidence for the solution of the problem is already present.

Finally, no memory of previous instances is involved. The subject keeps a record of successive instances, the reasons why he generated them and their outcome. The introduction of memory into concept attainment studies is an interesting, but essentially gratuitous variable which may distort inductive reasoning by placing an additional burden on the subject.

FIGURE 1

Numbers	Reasons for choice	Conforms	Does not conform
2 4 6		√	

Record sheet

PROCEDURE

The subjects, 29 psychology undergraduates (17 men and 12 women), were examined individually and instructed as follows:

"You will be given three numbers which conform to a simple rule that I have in mind. This rule is concerned with a relation between any three numbers and not with their absolute magnitude, i.e. it is not a rule like all numbers above (or below) 50, etc.

Your aim is to discover this rule by writing down sets of three numbers, together with reasons for your choice of them. After you have written down each set, I shall tell you whether your numbers conform to the rule or not, and you can make a note of this outcome on the record sheet provided. There is no time limit but you should try to discover this rule by citing the minimum sets of numbers.

Remember that your aim is not simply to find numbers which conform to the rule, but to discover the rule itself. When you feel highly confident that you have discovered it, *and not before*, you are to write it down and tell me what it is. Have you any questions?"

Subjects then wrote down their first set of numbers, under the numbers 2, 4, 6, on the record sheet (see Fig. 1), together with the reasons why they had chosen them. The Experimenter then said "those numbers do conform to rule," or "those numbers do not

conform to the rule," according to whether they were in increasing order of magnitude. The second set of numbers was then written down by the subject, the Experimenter giving the appropriate information as before. This procedure continued until the subject wrote down a rule. If the rule was the correct one, the experiment was concluded. If it was incorrect, the subject was told so and instructed to carry on as before. The experiment continued until the correct rule was announced, or the time for the session exceeded 45 minutes, or the subject expressed a wish to give up. The time was recorded and the implications of the results discussed. Finally, subjects were warned not to talk about the experiment.

Results

Quantitative

Results will be classified as follows:

(*a*) frequency of correct and incorrect rules,

(*b*) extent of enumerative and eliminative thinking,

(*c*) frequency of negative instances,

(*d*) immediate response to incorrect rules, and

(*e*) types of incorrect rule.

(*a*) *Frequency of correct and incorrect rules.* Table I shows the frequency of successive announcements of rules. The first announcement can either be correct

TABLE I

FREQUENCY OF ANNOUNCEMENTS

(n = 29)

First Announcement			*Second Announcement*			*Third Announcement*			*Fourth Announcement*			*Fifth Announcement*		
Immediate Correct	None	Incorrect	Correct	None	Incorrect	Correct	None	Incorrect	Correct	None	Incorrect	Correct	None	Incorrect
6	1	22	10	3	9	4	3	2	0	1	1	1	0	0

(defined as the "immediate correct announcement") or incorrect. (In nearly all cases incorrect rules were sufficient ones, e.g. "increasing intervals of two.") However, a third category, "none," is included for those subjects (one in this case) who made no announcement of a rule of any kind throughout the experiment. Those cases which fall into the "incorrect" category at the first announcement are redistributed within the second announcement categories, according to whether the second rule announced is correct or incorrect. But once again, the category "none" is included to cover those subjects who made no subsequent announcement of a rule of any kind after their first incorrect one. Similarly, those cases which fall into the second incorrect announcement category are redistributed within the third announcement categories on the same basis. Thus, the table provides a running record of the behaviour of the subjects.

In the sample there were 11 undergraduates in their first year, 12 in their second year and six in their third year. Of the six subjects who made the immediate

correct announcement, four were in their second year and two in their third year. There were no sex differences and no differences between Arts and Science background.

(*b*) *Extent of enumerative and eliminative thinking.* An index of the kind of thinking used was constructed by considering the nature of the instances in relation to the reasons given by the subject for choosing them. Each "reason for choice" was classified as either compatible, or incompatible, with all subsequent instances (including that instance for which each was specifically given). The total number of instances, compatible and incompatible with reasons, was computed for each subject, (i) up to the first rule announced, and independently (ii) from the first rule to the second one (if any). It was assumed that thinking was enumerative if there were a high proportion of compatible instances, and eliminative if there were a high proportion of incompatible ones.

The ratio of the number of incompatible, to compatible instances provided an eliminative/enumerative index for each subject. The mean ratio was 1·79 ($n = 6$) for those who made the immediate correct announcement, and 0·24 ($n = 22$) for those who made a first incorrect announcement. As predicted on the basis of the logical considerations given in the Introduction, a significant difference was obtained between these two means ($p = 0{\cdot}0002$, one tail test). The statistic used was Whitfield's (1947) extension of Kendall's S (1948) to a dichotomous variable.

The mean number of instances generated before making the immediate correct announcement was 8·0 (range 5 to 9), and the mean number before making a first incorrect announcement was 3·68 (range 1 to 7). This difference is, of course, highly significant. Thus, subjects who arrived at merely sufficient rules did so on the basis of relatively few confirming instances, while those who attained the necessary-and-sufficient rule tended to eliminate sufficient ones.

For those subjects whose *second* announcement was (i) correct, and (ii) incorrect, the mean ratios were 0·50 and 0·19 respectively. The difference between them was just short of statistical significance ($p = 0{\cdot}08$, one tail test).

It should be noted that in 18 out of the 22 cases of a first incorrect rule, the reason given for the subject's first instance was subsequently announced as at least a part of their first rule.

Figure 2 shows Vincent curves (Hilgard, 1938) of the mean number of incompatible instances (per fifths of the total number of instances for each subject) for (i) the immediate correct announcement, (ii) the first incorrect announcement, and (iii) the first incorrect announcement, omitting 15 subjects without any incompatible instances.

(Vincent curves represent the performance of subjects, equated in terms of their progress towards a criterion which is not defined by a fixed number of trials. The total number of trials taken by each subject to reach this criterion is divided into equal fractions, e.g. fifths, tenths, etc., and the progress made within each successive fraction is recorded.)

(*c*) *Frequency of negative instances.* The ratio of the number of negative instances of the correct rule to the total number of instances generated was computed for each subject, (i) up to the first rule announced, and (ii) from the first rule to the second one (if any). The mean ratio was 0·21 ($n = 6$) for those who made the immediate correct announcement and 0·04 ($n = 22$) for those who made a first incorrect announcement. As predicted, a significant difference was obtained between these two means ($p = 0{\cdot}0002$, one tail test).

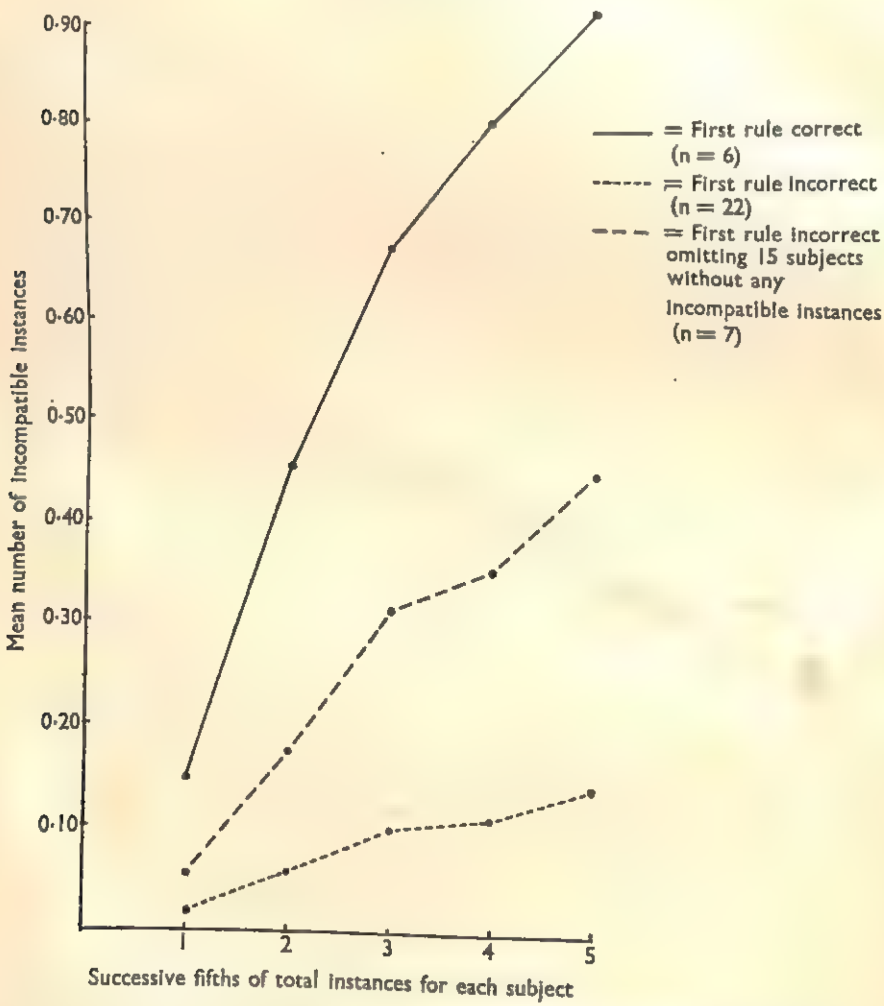
For those subjects whose *second* announcement was (i) correct, and (ii) incorrect, the mean ratios were 0·38 and 0·09 respectively, and the difference between them was significant ($p = 0{\cdot}003$, one tail test).

A highly significant correlation was obtained between the eliminative/enumerative index and the negative instance index, the value of tau being $+\ 0{\cdot}72$ ($p < 0{\cdot}00003$,

one tail test). It should be noted that a negative instance of the correct rule does not suffice to eliminate a sufficient rule. For example, a negative instance such as 6, 4, 2 is compatible with the sufficient rule, "intervals of two in increasing magnitude." A sufficient rule can only be eliminated by a positive instance which is incompatible with it.

(*d*) *Immediate response to incorrect rules.* The instance which is generated immediately after an incorrect announcement can either be compatible or incompatible with the rule announced, and it can be either a positive or a negative instance of the correct rule. Table II shows the frequency of cases in these four categories.

FIGURE 2



Vincent curves of the mean number of instances incompatible with reasons up to the first rule.

On a purely logical criterion it would be expected that, when subjects knew that their rule was incorrect (by being told so), they would depart from it and try a new one. But it will be seen from the table that in more than half the cases the rule is maintained, even though some other attribute, e.g. order, may be tested.

(*e*) *Types of incorrect rule.* Four different kinds of incorrect rule were announced most frequently: (i) numbers increasing in intervals of 2, i.e. a $(a + d)$, $(a + 2d)$ where $d = 2$; (ii) increasing multiples of the first number, i.e. a $(a + d)$, $(a + 2d)$ where $a = d$; (iii) consecutive even numbers, i.e. $2a$, $(2a + d)$, $(2a + 2d)$ where $d = 2$; and (iv) arithmetic progression, i.e. a, $(a + d)$, $(a + 2d)$.

In addition, five rules were announced only once: (i) "the first number added to the second number gives the third"; (ii) "numbers which add up to twelve"; (iii) "ascending progression formed by adding or multiplying by a constant"; (iv) "arithmetic or geometric progression"; and (v) "the second number is the first number plus one and the third is the first number plus four." All these rules, except the last one, are consistent with the initially given instance.

TABLE II

FREQUENCY OF TYPES OF INSTANCE IMMEDIATELY FOLLOWING INCORRECT RULES

	Compatible with incorrect rule	*Incompatible with incorrect rule*	*Total*
Positive instances	11	13	24
Negative instances	5	2	7
Total	16	15	31

Table III shows the frequency of incorrect rules at successive announcements. The five rules announced only once by five different subjects are classified together under "others." The frequency of correct rules is included for comparison. It will be noticed that "increasing intervals of two" is by far the most frequent at the first announcement but that there is only one case of it as a second announcement. On the other hand, there were twice as many cases of the more general rule, "arithmetic progression," at the second announcement than there were at the first. And this is obviously because the rules are not logically independent. "Consecutive even numbers" entails "increasing intervals of two," which in turn entails "arithmetic progression." Similarly, "multiples of the first number" entails "arithmetic progression" and, of course, all these rules entail the correct one, "increasing magnitude."

TABLE III

FREQUENCY OF INCORRECT RULES AT SUCCESSIVE ANNOUNCEMENTS

Rules	*Successive Announcements*				
	1st	*2nd*	*3rd*	*4th*	*5th*
Increasing intervals of two	8	1	0	0	0
Multiples of first number	4	1	0	0	0
Consecutive even numbers	3	0	0	0	0
Arithmetic progression	3	6	2	1	0
Others	4	1	0	0	0
Correct rule	6	10	4	0	1

Qualitative

Six protocols are given below, the first three from subjects who made the immediate correct announcement, and the second three from those who made incorrect announcements.

The most interesting qualitative feature of the results is that the successive announcements of three subjects consisted in the repetition of a single rule in different terms from those used on their previous announcements. In addition, three subjects reformulated their rules immediately after being told that they were incorrect, i.e. before generating any further instances. (This latter type of repetition was not counted as a new announcement in the quantitative results.)

In protocol No. 4 it will be seen that the second rule reads, "the middle number is the arithmetic mean of the other two," and the third reads, "the difference between two numbers next to each other is the same." These rules do not mean exactly the same thing. The first can be fulfilled by instances such as 5 7 9 and 9 7 5, the second by 5 7 9, 9 7 5 and 5 7 5. But, in any case, their connotations are wider than the instances given for them warrant. But the fourth rule, "adding a number, always the same one to form the next number," which follows the negative instance, 12 8 4, is restricted in scope so that it could not be fulfilled by some instances which could fulfil the two previous rules. Its connotation is now exactly that of arithmetic progression. It is not clear whether this subject, (i) appreciated the fine difference between the second and third rules, in spite of the fact that their instances conformed to arithmetic progression, or (ii) thought that there were no differences between the rules, but assumed that their expression was relevant, or (iii) thought that the rules were completely different, failing to realize that arithmetic progression was their common factor.

The following is a short extract from protocol No. 5.

> The rule is that the central figure is the mean of the two external ones.
>
> 6 10 14: the difference between the first two numbers, added to the second number gives the third; 7 11 15: to test this theory; 2 25 48: to test this theory.
>
> The rule is that the difference between the first two figures added to the second figure gives the third.

Thus, in this case it appears from the words, "to test this theory," as if the subject was announcing what she took to be a new rule, rather than merely reformulating the previous one.

The third example is found with the rule of successive multiples.

> The rule is to start with a basic number, then double it and thirdly to multiply it by three.
>
> 14 28 42; 7 14 21: the first number being a half of the second and a third of the third; 8 16 24: same reason; 9 18 27; 50 100 150.
>
> The rule is that the second number is double the first and two-thirds of the third.

It seems likely that in these cases the subjects cannot change their concepts (as judged by their instances) but change their description of them. Changing the description of a rule makes them think they have changed the rule itself: a verbal adjustment is made to satisfy the demand for a different hypothesis. These observations seem to be related to the findings of Hull (1920) and Smoke (1932) that a concept can be accurately employed before it can be correctly verbalized. Both results testify to the fact that insight into the mental processes involved in thinking is often defective. The present phenomenon is also reminiscent of magical thinking. The failure of a spell or curse can always be ascribed to some inexactitude in its utterance, rather than to its intrinsic deficiency: it would have worked, if only it had been pronounced correctly. The immediate repetition of the second rule in protocol No. 6 is particularly striking in this context.

The rule is that the three numbers must be in an ascending series and separated by regular step intervals.

The rule is that the first number can be arbitrarily chosen; the second number must be greater than the first and can be arbitrarily chosen; the third number is larger than the second by the same amount as the second is larger than the first.

Examples of the Protocols

The words which follow an instance are its "reason for choice." Negative instances are printed in italics.

A. *Immediate correct announcement*

No. 1. *Female, aged 25, 3rd year undergraduate*

12 24 36: unit figures are even and increase in two's; 8 10 12: even numbers increasing in two's; 2 6 10: even numbers increasing in four's; *6 4 2:* even numbers decreasing in two's; 2 6 8: even numbers ascending; 8 54 98: even numbers ascending; 1 17 23: ascending numbers; 1 18 23: ascending numbers; 1 2 3: ascending numbers.

The rule is ascending numbers (9 minutes).

No. 2. *Female, aged 21, 2nd year undergraduate*

3 6 9: three goes into the second figure twice and into the third figure three times; 2 4 8: perhaps the figures have to have an L.C.D.; 2 4 10: same reason; 2 5 10: the second number does not have to be divided by the first one; *10 6 4:* the highest number must go last; 4 6 10: the first number must be the lowest; 2 3 5: it is only the order that counts; 4 5 6: same reason; 1 7 13: same reason.

The rule is that the figures must be in numerical order (16 minutes).

No. 3. *Male, aged 25, 2nd year undergraduate*

8 10 12: continuous series of even numbers; 14 16 18: continuous series of even numbers; 20 22 24: continuous series of even numbers; 3 5 7: continuous series of odd numbers; 1 2 3: continuous series but with smaller intervals; *3 2 1:* reverse; 2 4 8: doubling series; *2 2 4:* two numbers the same; *6 4 2:* reverse of original numbers; 1 9 112: simple ascending numbers.

The rule is any ascending series of different numbers. (10 minutes).

B. *Incorrect announcements*

No. 4. *Female, aged 19, 1st year undergraduate*

8 10 12: two added each time; 14 16 18: even numbers in order of magnitude; 20 22 24: same reason; 1 3 5: two added to preceding number.

The rule is that by starting with any number two is added each time to form the next number.

2 6 10: middle number is the arithmetic mean of the other two; 1 50 99: same reason.

The rule is that the middle number is the arithmetic mean of the other two.

3 10 17: same number, seven, added each time; 0 3 6: three added each time.

The rule is that the difference between two numbers next to each other is the same.

12 8 4: the same number is subtracted each time to form the next number.

The rule is adding a number, always the same one to form the next number.

1 4 9: any three numbers in order of magnitude.

The rule is any three numbers in order of magnitude. (17 minutes.)

No. 5. *Female, aged 19, 1st year undergraduate*

1 3 5: add two to each number to give the following one; 16 18 20; to test the theory that it is simply a progression of two. These are chosen so that they are more complex and not merely simple numbers; 99 101 103: to test the progression of two theory, using odd numbers.

As these numbers can hardly have any other connection, unless it is very remote, the rule is a progression of adding two, in other words either all even or all odd numbers.

1 5 9: the average of the two numbers on the outside is the number between them.

The rule is that the central figure is the mean of the two external ones.

6 10 14: the difference between the first two numbers, added to the second number gives the third; 7 11 15; to test this theory; 2 25 48: to test this theory.

The rule is that the difference between the first two figures added to the second figure gives the third.

7 9 11, 11 12 13, *12 9 8, 77 75 71.*

Subject gives up. (45 minutes.)

No. 6. *Male, aged 23, 2nd year undergraduate*

8 10 12: step interval of two; 7 9 11: with numbers not divisible by two; 1 3 5: to see if rule may apply to numbers starting at two and upwards; *3 5 1:* the numbers do not necessarily have to be in ascending or descending order; *5 3 1:* could be in descending order.

The rule is that the three numbers must be in ascending order separated by intervals of two.

11 13 15: must have one number below ten in the series; 1 6 11: ascending series with regular step interval.

The rule is that the three numbers must be in an ascending series and separated by regular step intervals.

The rule is that the first number can be arbitrarily chosen; the second number must be greater than the first and can be arbitrarily chosen; the third number is larger than the second by the same amount as the second is larger than the first.

1 3 13: any three numbers in ascending order.

The rule is that the three numbers need have no relationship with each other, except that the second is larger than the first, and the third larger than the second. (38 minutes.)

Discussion

Only six of the 29 subjects gave the correct rule at their first announcement. These subjects tended both to eliminate more possibilities, and to generate more negative instances than did those who announced a first incorrect rule. Significant differences were obtained between these two groups on both criteria at the 0·0002 level of confidence.

On the other hand, the 13 subjects, who announced only one incorrect rule, presumably did so on the basis of simple enumeration, i.e. they assumed that confirming evidence alone justified their conclusions. But there are two other possible explanations of their behaviour.

Firstly, these subjects may have announced a rule in the hope that doing so would remove them from the experimental situation. This seems unlikely, however, because they all appeared highly motivated, and they expressed considerable surprise when told that their rule was not the one which the experimenter had in mind.

Secondly, they may have assumed that there could be only one rule to which the initially given instance could conform. Familiarity with the "number series" type of problem in which there is supposed to be only one correct continuation, might have induced a set for the one "right" answer. It could also be argued that the correct rule (increasing magnitude) was so trivial that students would have been reluctant to entertain it. However, the point is not that most subjects failed to give the correct rule at their first announcement, but that they adopted a strategy which tended to preclude its attainment.

But after they had announced one incorrect rule, 10 of these 13 subjects (three made no further announcement) gave the correct rule at their next announcement. And the results suggest that they did this by eliminating more possibilities, and generating more negative instances than those subjects who announced a second incorrect rule. The difference between these two groups was just short of significance ($p = 0{\cdot}08$) on the eliminative/enumerative index, and was significant at the 0·003 level on the negative instance index. Thus, it is possible that, at the beginning of the experiment, the reinforcement of these subjects' rules by their confirming instances blocked the notion that there might be any alternative. But after this set

had been broken by an incorrect announcement, they were able to eliminate any remaining alternative which occurred to them.

These possibilities can hardly apply to those nine subjects who announced two or more incorrect rules. For after their first announcement they had evidence to show that stating rules would not remove them from the situation, and that there could be more than one possible rule. Hence, it appears as if from then on they were reasoning by simple enumeration, and announcing sufficient rules, either from an inability to do otherwise, or from a preference for what Bruner has called a "direct test." In other words, they might not have known how to attempt to falsify a rule by themselves; or they might have known how to do it, but still found it simpler, more certain or more reassuring to get a straight answer from the experimenter about the correctness of their rules. This second possibility, however, seems rather remote because the method of elimination is not difficult to apply. The attempted falsification of a rule in no way depends on discovering a suitable alternative to substitute for it. All that the subject had to do is to generate an instance which is similar to previous positive ones, but does not conform to the tested rule. If the outcome is positive, the rule can be decisively eliminated. Thus, the subject has to reason that a rule will be false, if it does not cover a positive instance of the correct rule. He must, at some stage, relinquish a rule which may have been confirmed, and adopt that eliminative strategy which Bruner calls "conservative focussing" and which Mill called the method of difference.

The announcement of a sufficient rule is, in fact, the frequent result of enumerative thinking. In the present investigation, in contrast to Bruner's experiment, the use of confirming evidence alone will compel the announcement of a rule, as the only way of finding out whether or not it is the correct one. Here there are no ready-made instances displayed which might be used to correct a sufficient rule. On the contrary, instances exemplifying such a rule can never be exhausted. Thus, the experiment demonstrates the dangers of induction by simple enumeration as a means of discovering truth. In real life there is no authority to pronounce judgement on inferences: the inferences can only be checked against the evidence.

The results show that very few intelligent young adults spontaneously test their beliefs in a situation which does not appear to be of a "scientific" nature. The task simulates a miniature scientific problem, in which the variables are unknown, and in which evidence has to be systematically adduced to refute or support hypotheses. Generating an instance corresponds to doing an experiment, knowledge that the instance conforms, or does not conform, corresponds to its result, and an incorrect announcement corresponds to an inference from uncontrolled data. The kind of attitude which this task demands is that implicit in the formal analysis of scientific procedure proposed by Popper (1959). It consists in a willingness to attempt to falsify hypotheses, and thus to test those intuitive ideas which so often carry the feeling of certitude. The methodological analogue of this attitude consists in the use of increasingly stringent controls. Obviously scientific method can be taught and cultivated. But the readiness (as opposed to the capacity) to think and argue rationally in an unsystematized area of knowledge is presumably related to other factors besides intelligence, in so far as it implies a disposition to refute, rather than vindicate assertions, and to tolerate the disenchantment of negative instances. And certainly these qualities are no less important for thinking in general than the more obvious cognitive functions associated with purely deductive reasoning.

I am primarily indebted to Dr. A. R. Jonckheere. Frequent arguments with him about the logical issues involved in this research have greatly helped to clarify my ideas, and his constructive criticism of the manuscript has been invaluable. I should also like

to thank Professor G. C. Drew, Dr. R. P. Kelvin and Mr. R. D. Shepherd for valuable comments. Finally, I am indebted to my subjects for the great interest and enthusiasm which they expressed in the experiment.

References

Bruner, J. S., Goodnow, J. J., and Austin, G. A. (1956) *A Study of Thinking*. London.

Goldstein, K., and Scheerer, M. (1941). Abstract and concrete behaviour; an experimental study with special tests. *Psychol. Monogr.*, **53**, No. 239.

Hilgard, E. R. (1938). A summary and evaluation of alternative procedures for the construction of Vincent curves. *Psychol. Bull.*, **35**, 282–97.

Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychol. Monogr.*, **28**, No. 123.

Kendall, M. G. (1948). *Rank Correlation Methods*. London.

Popper, K. R. (1959). *The Logic of Scientific Discovery*. London.

Smoke, K. L. (1932). An objective study of concept formation. *Psychol. Monogr.*, **42**, No. 191.

Von Wright, G. H. (1951). *A Treatise on Probability and Induction*. London.

Whitfield, J. W. (1947). Rank correlation between two variables, one of which is ranked, the other dichotomous. *Biometrika*, **34**, 292.

EXPERIMENTAL DISORIENTATION AND CONCEPTUAL CONFUSION

BY

F. H. GEORGE and S. B. McINTOSH

From the Department of Psychology, University of Bristol

Thirty-two subjects were examined on a visual matching task. They were tested for their ability to maintain an orientation with respect to a particular direction in the horizontal plane, while being kept in circumstances designed to minimize their input of information and create thereby some conceptual confusion.

The results suggest that subjects tend to make corrections as if they were in the same position in space throughout, even though they have no necessary reason for supposing this to be true and some reasons for supposing the opposite. It seemed that non-verbal information had to be presented to the subjects in order to suppress this tendency.

Voluntary rotation of the subject from one setting to the next produced no more than chance errors, while arbitrary rotations only produced errors when cues were inconsistent, or possibly where no cues were available at all. In the cases where no cues—or minimal cues—were available, assumptions were made by the subjects about the nature of the environment. This caused both errors and correct responses for reasons that were not justified by the evidence in the hands of the subject.

Introspective reports revealed some interesting results as to the cues utilized and concepts consciously formulated by the subjects for their use.

INTRODUCTION

Since this experiment was an attempt, in part, to build upon some previous experiments by M. F. Piercy (1957) we shall briefly summarize his experiments and their results.

Piercy carried out an experiment in which subjects were confronted by a luminous arrow, pointing in one of the following directions: north, south, east, west, north-east, north-west, south-east and south-west. The subjects were then asked to perform a voluntary rotation through 180° in the darkened room and were asked to place another luminous arrow that they were then confronted by, in the *same direction in space* as the first arrow. This experiment was also carried out with some variations, such as in the use of "landmarks," but the principal results were that the subjects suffered from some considerable degree of confusion. Piercy also found that there was more confusion with respect to the fore and aft orientation of the arrow than with respect to the left-right orientation and this was thought by him to be attributable to the fact that there were two dimensions for error in the fore and aft direction and only one dimension in the left-right direction.

During an attempt to reproduce Piercy's results, which he described as "a preliminary exploration of factors relevant to the constancy of the perception of direction in small scale spatial orientation," it was found impossible to obtain any errors with normal subjects under the original conditions used by him. This was the case even when preliminary practice in a lighted room—part of Piercy's procedure—was omitted. However, there were certain differences between the procedure used here and that used by Piercy, and these we shall discuss presently.

As only five of the 28 subjects used by Piercy were "free from neurological and psychiatric disorder," while 16 were neurotic patients and a further seven were suffering from cerebral lesions, it seems reasonable to suppose that his results are only possible when the group of subjects used are retarded in the ability to form concepts. It is true that Piercy's five "normal" subjects showed no statistical differences from the "abnormal" ones, but even so they may have been of relatively low intellectual standards.

By the same token the subjects used in our experiments were all undergraduates and were generally of superior intelligence, at any rate superior to the "dull normal or above" group used by Piercy. The ease with which they performed the Piercy type of task was disconcerting, but this led us to consider the degree of confusion which it was necessary to introduce before the tasks showed systematic errors and thus to try and isolate the important factors in spatial orientation.

Now we must discuss briefly the way in which our own experiment differed from Piercy's.

As far as the results given in this paper are concerned, there was no subject who was treated in exactly the same way as any of the subjects used by Piercy. However the first condition was very similar. This first condition involved a voluntary rotation of 180° from a position inspecting one arrow to a new position inspecting another arrow. In fact, Piercy's conditions were at first used, and with some half a dozen subjects no faults were found to occur. In these preliminary experiments (many others were tried informally) the initial position of the variable arrow was kept constant, although subsequently, in the results given here, this constancy was not maintained.

Apart from these two differences, there were no other major differences between our procedure and Piercy's for the first condition, except that the blackout conditions in our case were apparently more complete. Nevertheless there will be a doubt as to how much a direct comparison of the two sets of results are possible, and we should like to lay our emphasis on these experiments as carrying Piercy's work one step further forward rather than conflicting with it. In other words a direct comparison with Piercy's results is not the major consideration in these experiments. What is quite central to our purpose is to try and discover the method by which *concepts* are formed.

It should be said straight away that our experiments were performed in the dark; the luminous arrow was all the subject was able to see and this was painted on a disc that was laid flat on a table in front of him. At the last stage of each trial the subject has to move the disc with his hands and place it in the new position as directed.

Experimental Procedure

Four conditions were used, the apparatus for each consisting of two tables approximately 2 ft. 6 in. tall and 4 ft. apart. On each table was a circular disc 1 ft. in diameter with a 6 in. luminous arrow on it.

Figures 1 and 2 show the eight arrow positions that were used. Figure 1 shows the initial orientations of the arrow relative to the observer and Figure 2 shows the equivalent orientations which were presented to him and from which he had to return the arrow to its initial orientation. Where two arrows were used, Figure 1 represents the fixed arrow and Figure 2 the variable arrow. The two tables carrying the discs were placed in a blacked out room and parallel to the walls. Approximately 3 seconds were allowed for the subject to inspect the luminous arrow, although this was difficult to control exactly owing to the room being blacked out. A swivel chair was used for conditions 2, 3 and 4. The four conditions were:—

Condition 1

A repetition of condition 1 of the experiment carried out by Piercy, with the exceptions already noted above, i.e. the subject viewed the standard arrow and then made a voluntary movement through 180° to adjust the variable arrow. No previous practice in the light was given.

Condition 2

The same arrow was used both as the standard and variable, so that the subject was facing the same direction all the time but was rotated three or four times between the presentation of the standard and variable arrow. The subject had no means of knowing

that he was to finish his rotation facing the same direction as he started, indeed he was told that he would then see "another arrow," implying that there were two distinct and separate arrows in the experiment.

Condition 3

This condition was identical with the conditions of 2 except that the subject was further confused by the removal of at least one other possible cue. The experimenter *stressed* more strongly than in condition 2 the fact that the variable arrow might not be in the same

FIGURE 1

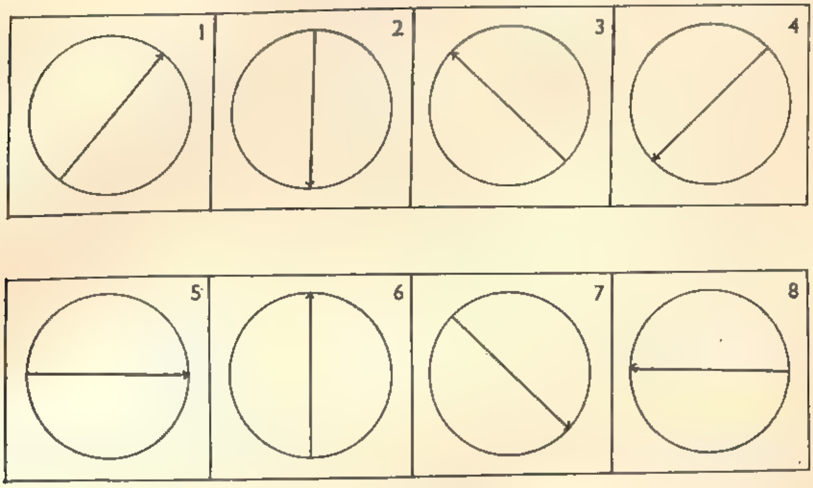
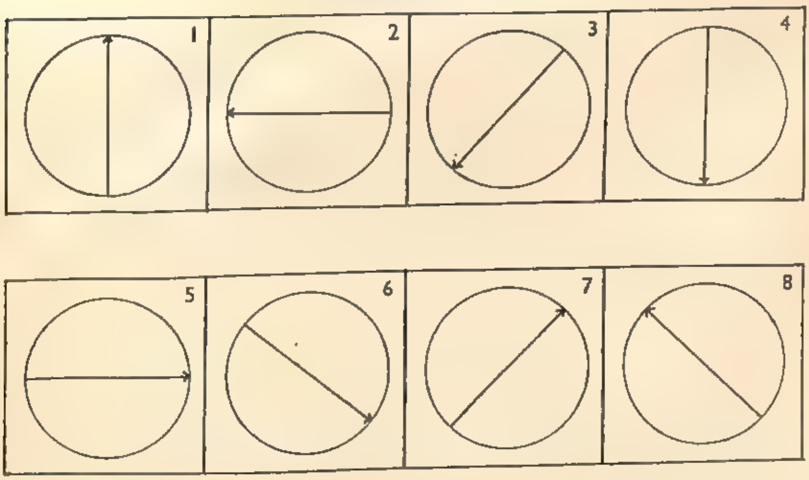


FIGURE 2



position in space, and, more important, the experimenter himself changed his position during rotation so that his position could not be used as a cue (a fixed point in the environment).

Condition 4

In half the trials two arrows were used and the other half only one in random order, so that sometimes the subject was facing in the same direction to adjust the variable arrow and sometimes in the opposite direction. The experimenter changed position as in

condition 3 while rotating the subject and the emphasis on the possibility of two arrows was only as strong as in condition 2 since there really were two arrows now in half the trials.

No knowledge of results was given. All the subjects were given preliminary instructions, before entering the dark room, as follows:—

All conditions

"I am going to take you into a dark room and I want you to keep your eyes shut unless I specifically ask you to open them. (This is because I shall have to put the light on from time to time and if you open your eyes they will become light adapted and you will find the experiment more difficult). It was explained that the arrows were on tables in front of them and they were told not to touch them until they made their final adjustment to the arrow.

I will ask you to open your eyes and you will find in front of you a two-dimensional luminous arrow. I want you to note the direction in which the arrow is pointing in space, not in relation to yourself, but in relation to space, i.e. whether it is pointing towards Birmingham or Timbuctoo." Now the four conditions differ slightly.

Condition 1:—

"I then want you to turn round, and behind you there will be a second arrow. I want you, without looking back, to point this arrow in the same direction in space as the first arrow."

Conditions 2 and 4:—

"I will then rotate you and show you *another* arrow which I want you to point in the same direction in space as the first arrow."

Condition 3:—

"I will then rotate you and show you *another* arrow which will not necessarily be in the same position in the room as the first arrow."

It was made quite certain that the subject understood what he was expected to do, before he was taken into the dark room. The instructions were then repeated as the subject performed each step.

For conditions 2, 3 and 4 the procedure was as follows:—

"Open your eyes. Note the direction in which the arrow is pointing in space—Close your eyes."

When the same arrow was being used as the variable it was then adjusted by the experimenter but the light was always put on at this juncture to avoid the turning of the light on or off being a cue.) The subject was then rotated and instructed to "adjust the other arrow."

Each subject was tested with eight intercardinal directions (Figs. 1 and 2). Responses were recorded on a record sheet and at the end of the session each subject was asked:—

(i) "On what basis did you make your judgements?"

(ii) "Did you use anything as cues?"

Subjects

Thirty-two undergraduates of the University of Bristol were tested, eight in each group. No attempt was made to measure the intelligence of the subjects, but it was assumed that they were more intelligent than the subjects used by Piercy whom he described as being of "dull normal intelligence or above."

Results

Piercy's method of ignoring all errors of 10° or less was followed in drawing up Table I.

Condition 1. Only two errors occurred. These were both at the beginning of a series and could be explained on the grounds that the subject had not, in fact, fully understood what was expected of him, as no preliminary practice was given.

Condition 2. Only four errors occurred. In three of these cases they were the last responses in the series and were duplications of the position of the variable arrow. Both these factors suggest there may have been boredom or fatigue at this stage resulting not in disorientation but in poor memory so that the first position of the arrow was forgotten and only the second remained as a guide.

TABLE I

Position	*Condition 1*								*Condition 2*							
1	−	+	+	+	+	+	+	−	+	+	+	+	+	+	+	+
2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
5	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
6	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+
7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
8	+	+	+	+	+	+	+	+	−	−	+	+	−	+	+	+
Errors	1	0	0	0	0	0	0	1	1	2	0	0	1	0	0	0

Position	*Condition 3*								*Condition 4*							
1	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+
2	+	−	−	+	−	+	+	+	+	+	+	+	+	+	+	+
3	−	+	−	+	−	−	+	+	−	−	−	−	−	+	−	−
4	−	+	+	+	+	−	−	+	−	+	−	+	−	−	+	+
5	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−
6	−	+	+	+	−	−	+	+	+	−	−	−	−	−	−	−
7	+	+	+	−	+	[illegible]	+	+	−	−	−	−	−	−	−	−
8	−	+	+	+	−	−	+	+	+	+	+	−	−	+	−	+
Errors	4	1	2	1	4	4	1	0	5	4	5	5	6	4	5	4

Each column shows the eight different performances of the same subject.
+ means a success (error less than 10°).
− means a failure (error more than 10°).

Condition 3. Only one subject successfully completed the series. Seventeen errors occurred in all, only three of these being duplications of the position of the variable arrow, none of them occurring as a final response in a series. Eight errors occurred in the cardinal directions, three in the left-right and five in the fore-aft dimensions.

Condition 4. No subject successfully completed the series. A total of 38 errors occurred, none being duplications of the position of the variable arrow. Twenty-one errors were in both dimensions.

With respect to cardinal errors in this last condition, there were 10 in the left-right and six in the fore-aft dimension. Introspective reports were of some interest. Table II shows these results in tabular form.

We now show a brief statistical analysis of these results where the statistical difference between the means of the groups under different conditions was calculated for comparison purposes. The table shows the comparison between groups under conditions 1 and 4; 1 and 3; and 3 and 4. In Table III the probabilities were computed for a fiducial level of 0·05.

TABLE II

Cues used (conceptual and otherwise)	*Condition* 1	*Condition* 2	*Condition* 3	*Condition* 4
1. The subject's own person	0	5	3	3
2. A clock face	1	1	0	3
3. A compass face	0	1	1	0
4. A count made (or attempted) of number of times rotated	0	1	0	2
5. The position of the experimenter in the room	0	0	2	3
6. Parallel lines and imagined geometrical frame-works	2	0	0	0
7. The shape of the room implied by the table which was felt at the beginning and assumed in alignment with room	4	0	0	0

This table shows a list of cues used in the experiments.

TABLE III

Comparison	*t- valued obtained*	*t- value from t- tables*
1 and 4	$t = 4{\cdot}3$	$n = 16, t = 2{\cdot}145$
1 and 3	$t = 1{\cdot}4$	$n = 16, t = 2{\cdot}145$
3 and 4	$t = 3{\cdot}9$	$n = 16, t = 2{\cdot}145$

This table shows difference between means with respect to the most important groups classified according to their experimental treatment.

Discussion

The main conclusion that one can draw from these results seems to be that reduction of visual cues has to be virtually complete before disorientation begins to occur in normal subjects. The findings of Asch and Witkin (1948) which were to the effect that visual cues are used in preference to tactile, proprioceptive and vestibular cues when there is conflicting evidence, do not rule out the possibility of these other cues being adequate for orientation.

Even when auditory and other cues are also removed little disorientation occurs unless the subject's *set* is inappropriate. This apparently occurs in the Ames' experiments where the set is towards a "normal environment."

It might be worth trying to measure the strength of this expectancy by preceding conditions 2, 3 or 4 by condition 1 so that the subject's immediately preceding experience might then bias him to expect a "180° environment."

From the introspective reports, it seems clear that subjects built up a sort of conceptual framework upon which they acted, even though some of the cues they had been verbally given conflicted with it. Thus subjects who were told that the second arrow might not be in the same position in space as the first still reported having the idea of a clock, or compass, etc., and using themselves as a cue, and regarding themselves as having exactly the same relation to the second arrow as the first. When the inconsistency of this was pointed out to them afterwards, subjects either said they had forgotten this information or became slightly annoyed asserting they could not have acted rationally at all, had they not *assumed* some fixed position.

In the latter case there seems to be an analogy with experimental neurosis in animals, though as there was no knowledge of results (reward or punishment) at the time, frustration did not occur until afterwards.

During preliminary investigations and in condition 2 it was found that many factors other than visual cues were utilised in orientation. The position in which the experimenter was standing, the number of times the subject was rotated are examples of such cues that were used. These additional cues perhaps explain the comparative lack of errors.

In condition 3 an attempt was made both to eliminate these cues and produce confusion by adding the statement that the subjects would be shown "another arrow which will not necessarily be in the same position in space as the first arrow." Condition 4 was similar without quite the same emphasis on the factor of a second arrow. It might have been interesting to eliminate the cues without introducing confusion and see whether confusion would arise spontaneously.

The very significant difference in errors between conditions 1 and 4 may have been due to the total absence of guiding cues and the disruption of normal concept formation by making cues misleading, e.g. the position of the experimenter did not remain constant.

The fact that the majority of the errors in condition 4 (20/39) occurred when the second arrow was placed at 180° from the correct direction and would thus have been correct had the arrow been in the *original* position suggests a relatively consistent attitude on the part of the subject.

If the hypothesis that concept formation is a primary factor in orientation is correct, this could account for the difference between our results and Piercy's results, as our subjects should be superior at "concept formation." It might also account for the fact that we found no significant difference between errors in the fore-aft dimension and those in the left-right dimension. Piercy ascribed this to "contamination between the two systems of conceptual analysis which identify the vertical and fore-aft dimensions of visual space," and this could be overcome with superior skill at forming new concepts, although here the difference in our conditions may be all-important.

We shall try, very briefly, to interpret the results of conceptual confusion in a cybernetic manner. We should, for example, guess that items were occurring in the input system which did not normally need discriminating (were generalized) but which now needed such discrimination. This seems to be possible both in the identification of objects and in the identification of relations, and in these experiments it is the latter which are important. In general, we have in the first case two events represented by collections of property names such as *abcdefgh* or *abcdefghi* which are normally identified and which now have to be distinguished. In the second case we have a relation Rab and we are simply disguising the characteristics $r, s, \ldots, w$ which define R, i.e.:—

$$R = f(r, s, \ldots, w)$$

for some function f.

Such results as these should be capable of being integrated into a hardware model or a blueprint for a hardware model (an effective theory) although no detailed attempt will be made to do so here.

We should, though, try and give some account of the significance of our notion of concept formation, and we can most appropriately describe it in terms of our briefly mentioned cybernetic model. This is not meant to imply that other interpretations are impossible, only that this seems a convenient manner in which to interpret the present results.

By concept formation we shall mean the setting up of functions f, g, . . . in the memory store such that with recognition (Uttley, 1954, 1955; George, 1956, 1957, 1959; Culbertson, 1949) a comparison takes place between the input event and the events in store. If a sequence of symbols representing an event is incomplete, as it will be in general, then the problem of recognition is to ascribe a probability to the subset being a member of a particular complete set; that complete set which we call the event (or event name) itself. Now there will be many circumstances met by our model (it hardly matters whether we think of this as a computer or an organism) which will require that a decision be made as to whether the event is a new event or not. If not, we shall say that stimulus generalization has taken place, otherwise a new function is formed and we shall call this "concept formation." The conditions for the decision are the same as those for a discrimination in a computer programme. The crucial variable here is probably dependent upon the strength of the probability, how many characteristics of the complete set seem to be missing from the subset, how many incorrect ones are included in the observed subset, and the "value" (motivational importance) (George, 1957) of the expected next response and subsequent stimulus state.

Now in the conceptually confusing situation that was met with in Piercy's and our experiments, the implication is that unusual discriminations were called for. For example, it is not important pragmatically, in the normal course of events, to distinguish the spatial frame of reference in two presentations of an arrow in a room either because the frames are obvious, visually, or because they are to be acted upon, if at all, relative to the observer rather than the spatial frame. Thus in our experiments it is the relational data of the arrow with respect to its spatial frame of reference that have been destroyed, or at least disguised. As a result there was a tendency to identify two different events (stimulus generalization) or to form new concepts, and although these were sometimes arbitrary and inadequate they were in the easier cases more effective than the method used by those subjects who tended not to be able to form new concepts at all. Thus the contamination that Piercy observed was made possible because of the lack of this ability to discriminate two different sets of relations, i.e. in our terminology, to form new concepts when needed.

References

Asch, S. E., and Witkin, H. A. (1948). Studies in space orientation: I and II. *J. exp. Psychol.*, **38**, 325–37; 455–77.

Culbertson, J. T. (1949). *Consciousness and Behavior.* Iowa.

George, F. H. (1956). Logical networks and behaviour. *Bull. Math. Biophysics*, **18**, 337–48.

George, F. H. (1957). Logical networks and probability. *Bull. Math. Biophysics*, **19**, 187–99.

George, F. H. (1959). Models and theories in social psychology. In Llewellyn Gross (Ed.) *Symposium on Sociological Theory.* New York.

Piercy, Malcolm (1957). Experimental disorientation in the horizontal plane. *Quart. J. exp. Psychol.*, **9**, 65–77.

Uttley, A. M. (1954). The classification of signals in the central nervous system. *Electroencph. clin. Neurophysiol.*, **6**, 479–94.

Uttley, A. M. (1955). The conditional probability of signals in the nervous system. *RRE-Memorandum No.* 1109.

REDUNDANCY AS AN EXPERIMENTAL VARIABLE

BY

ALAN STANILAND

From the Psychological Laboratory, University of Exeter

This paper is concerned with the measure of redundancy derived from Information Theory, as an "objective" specification of material used in perceptual experiments. As examples for discussion it cites especially experiments by Fitts and his associates (1956), on redundancy in visual patterns, and by Miller (1958) on recall of redundant strings of letters. It offers no suggestion or comment on the use of the redundancy concept in neurophysiological models, nor does it question the concept as applied to ordinary language. Language has a long-term background, against which particular experiments (or cross-word solving or word-guessing) occur: efficient guessing or utilization of probabilities can be referred to well established habits. In reduced situations, such as laboratory experiments with artificial material, this is not the case, and it will be argued that redundancy as calculated in the above experiments does not stand for a constituent of the experimental situation which can be assessed *a priori* and treated as an experimental variable comparable with a physical measure such as length. Finally, a different method of calculation, based on an alternative view of redundancy, will be put forward and related to the results of the experiments discussed.

A brief description of redundancy may be taken from Miller (1951):

> "By favouring some sequences of verbal units more than others, we effectively reduce the amount of information that is encoded per unit. . . . Suppose the various possible messages in a language consist of sequences of three digits: '1–9–3' or '8–7–0,' etc. So long as each digit in the message is unrelated to the other two, we can send any one of 10^3, or 1000 different messages. If the symbols are not independent, however, but if some fixed relation exists among them, the number of alternatives is reduced . . . our freedom of choice is restricted and the message cannot convey as much information."

Such restrictions, or constraints, on theoretical freedom have been shown to have several interrelated effects (c.f. Quastler, 1955):

(i) The actual uncertainty of the message source is reduced compared with the theoretical uncertainty. If constraints permit only 100 messages to be sent, the source uncertainty is $\log_2 100$ (6·64 bits) instead of $\log_2 1000$ (9·97 bits).

(ii) The coding of messages becomes uneconomical, e.g. 100 messages could be sent by using sequences of only two digits, giving 10^2 messages. Thus redundancy can be expressed mathematically as a relationship between two uncertainties, one for the theoretical range of messages, the other for the permitted number within the constraints. This is commonly reduced to percentage redundancy:

$$\frac{\text{H (theoretical source)} - \text{H (permitted selection)}}{\text{H (theoretical source)}} \times 100$$

In the case above:

$$\frac{\log_2 10^3 - \log_2 10^2}{\log_2 10^3} \times 100 = \frac{9{\cdot}97 - 6{\cdot}64}{9{\cdot}97} \times 100 = 33{\cdot}3 \text{ per cent.}$$

(iii) Because some sequences are excluded, internal or transition probabilities greater than 1/10 (for 10 digits) appear between digits. If, for example, 5 is always followed by 6 or 7, when 5 appears there are only two possibilities open for the next digit: p (6) = p (7) = 1/2, giving an uncertainty of 1 bit. By assessment of these transition uncertainties over a representative sequence of messages we can infer characteristics of the source without direct knowledge of it.

Production of Redundant Material

In each of the experiments cited, two sets of material were produced, one of which was called "random," the other "redundant" with specified percentage redundancy assigned to it. In each, comparison was made between results obtained from the

two kinds of material. We shall consider the methods of producing and assigning redundancy to the material, without at first discussing the experimental contexts.

Fitts's metric figures

In an experiment by Fitts *et al.* (1956), histogram-like patterns were used in a recognition task. They were of a type called by Fitts "metric figures," produced by filling cells in a 4×4 or 8×8 matrix. The 8×8 range were chiefly used, but only the 4×4 will be mentioned as examples, because their smaller number allows a briefer analysis. Each "bar-detail" (or upright) was composed of 1–4 filled cells, so that the range of possible patterns is $4^4 = 256$, equivalent to an 8 bits task. From this range Fitts took two selections, one random from all patterns, the other a non-random selection in which bars of 1, 2, 3, 4, appeared once and once only. The number of these is $4 \times 3 \times 2 \times 1 = 24$, equivalent to 4·58 bits. This selection was called "constrained" and supplied the redundant material, its redundancy being calculated as $\frac{8 - 4{\cdot}58}{8}\ 100 = 43{\cdot}5$ per cent.

The figure 4·58 expresses the average sum of transition uncertainties within the selected patterns, which can equally well be discussed as number groups. This is shown if we break down the constrained set into the successive uncertainties for each digit. Using A, B, C, D, for the four bar-details, any item then gives:

Selection 1 (constrained):	A	B	C	D	
Transition probability:	0·25	0·33	0·5	1·0	
Equiv. uncertainty:	2·0	1·58	1·0	0	= 4·58 bits

i.e. A can be chosen from four sizes: when one has been chosen, B is limited to the other three, C to two, and finally D is determined, there being only three degrees of freedom.

If we tabulate the uncertainties for each item (as is done for another selection in Table I) we obtain a matrix of 24 rows with each row identical, the average for each column therefore being the same as the uncertainties in the rows.

The mean value is not unique to this selection and another can be shown having the same sum:

Selection 2	A	B	C	D	
Bar height:	1	1 or 2	1–3	1–4	
Transition probability:	1·0	0·5	0·33	0·25	
Equiv. uncertainty:	0	1·0	1·58	2·0	= 4·58 bits.

The component uncertainties are the same though their order is different. Like the first selection this contains 24 patterns. Thus 4·58 as a measure applies to both selections and could make no distinction between them. It can further be shown that the same figure is obtainable as a mean value for *any* 24 patterns or groups drawn from the same source. The selection set out in Table 1 was taken from a random number table with a few alterations to provide illustrative points. It can be called the near-random selection. To employ the method of decomposition into successive choices given by Shannon (Shannon and Weaver, 1949), the selection is ordered in terms of its first digits, the uncertainty calculated at each choice, and the results summed by rows and columns and averaged. Thus:

For the first digit choice we have:

p (1) = 8/24	0·528
p (2) = 1/24	0·191
p (3) = 4/24	0·431
p (4) = 11/24	0·516
Σ-$p \log_2 p$	= 1·666

TABLE I

DECOMPOSITION OF SELECTION 3

No.	*Groups*	*A*	*B*	*C*	*D*	*Group totals*	*"Core"*
1	1123	1·67	1·5	1·0	0	4·17	112
2	1144	1·67	1·5	1·0	0	4·17	114
3	1211	1·67	1·5	1·0	0	4·17	121
4	1232	1·67	1·5	1·0	0	4·17	123
5	1421	1·67	1·5	0·81	1·58	5·56	1421
6	1422	1·67	1·5	0·81	1·58	5·56	1422
7	1423	1·67	1·5	0·81	1·58	5·56	1423
8	1433	1·67	1·5	0·81	0	3·98	143
9	2113	1·67	0	0	0	1·67	2
10	3341	1·67	0·81	0	1·58	4·06	3341
11	3342	1·67	0·81	0	1·58	4·06	3342
12	3343	1·67	0·81	0	1·58	4·06	3343
13	3421	1·67	0·81	0	0	2·48	34
14	4122	1·67	1·79	0·92	1·0	5·38	4122
15	4123	1·67	1·79	0·92	1·0	5·38	4123
16	4134	1·67	1·79	0·92	0	4·38	413
17	4224	1·67	1·79	1·0	0	4·46	422
18	4234	1·67	1·79	1·0	0	4·46	423
19	4331	1·67	1·79	0·72	2·0	6·18	4331
20	4332	1·67	1·79	0·72	2·0	6·18	4332
21	4333	1·67	1·79	0·72	2·0	6·18	4333
22	4334	1·67	1·79	0·72	2·0	6·18	4334
23	4344	1·67	1·79	0·72	0	4·18	434
24	4444	1·67	1·79	0	0	3·46	44
Totals:		40·08	34·93	15·60	19·48	110·09	
Means:		1·67	1·455	0·650	0·81	4·585 bits	

Note: The "core" leaves out terminal digits of 0 bits, and internal redundant digits are underlined.

Similarly, given 1 as first digit, the second digit probabilities are:

$$
\begin{array}{lll}
p(1) = 2/8 & & 0\cdot5 \\
p(2) = 2/8 & & 0\cdot5 \\
p(3) = 0 & & 0 \\
p(4) = 4/8 & & 0\cdot5 \\
\hline
\Sigma\text{-}p \log_2 p & & = 1\cdot5
\end{array}
$$

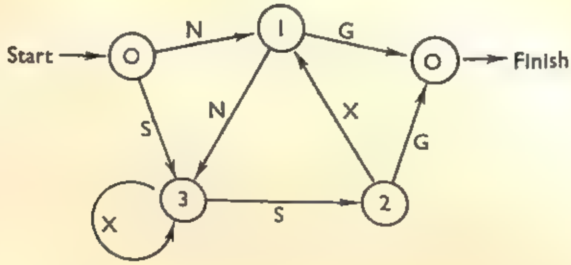
There are now three selections of the same number; each has the same average transition uncertainty; for all, calculation would give the same redundancy value relative to the source. Yet differences between them are obvious, and the question arises of how we can define the important ones. What, if anything, is indicated by their common average transition uncertainty of 4·58 bits?

Two differences stand out. In the first two selections all actual uncertainties are identical with the average uncertainties, but in the third the group sums vary between 1·67 and 6·18 bits, and all but the first column averages are different from any actual value in the column. Again, the first two selections have the uncertainty values

identically ordered, whereas in the third the rank order of uncertainties varies between groups. To these differences the average figure is insensitive.

We can provisionally express this difference by saying that the first two selections are "principled," while the third is "near-random." In the first two the principle might be called "simple" in the sense that it can be briefly stated, whereas any adequate description of Selection 3 must approach very closely to an enumeration of the contents. Other selections can be made more principled than Selection 3, but whose principle cannot be as briefly stated as those of Selections 1 and 2. Thus the evidence is for a gradation of differences, rather than an abrupt distinction.

FIGURE 1



Miller's Finite-State Generator.

We can now consider the similarities. Mathematically they are not surprising. Since the only datum contributed to the redundancy formula is $\log_2 24$ for a selection of 24 groups, it is not to be expected that the result will show any discrimination between selections of that number. What Shannon's decomposition shows is that uncertainty diminishes to zero at any point where no variation is open, and is exhausted as soon as each group is uniquely defined relative to *that* selection. Thus, in Selection 3:

Given first element/s: 2, group is No. 9. Group uncertainty, 1·67 bits.
44, group is No. 24. Group uncertainty, 3·46 bits.

But, 433, group may be Nos. 19–22. The group uncertainties are 6·18 bits, of which only 4·18 have so far been resolved.

The point may be put in another way. Fitts refers to the constrained selection as "non-random," and then calculates its redundancy from the fact that it consists of 24 items. The calculation, however, applies equally to any selection of the same number of items, and the fact that his selection happens to be "principled" (i.e. non random) is an independent fact which does not enter into the measure.

Miller's redundant strings

Another form of redundant material is used by Miller (1958) in memorization and recall tasks. The source was a finite-state generator (Fig. 1), equivalent to following permitted routes on a map with uni-directional constraints (indicated by arrows). By this method Miller obtained 22 "well-formed strings" of seven or fewer letters.

> "Of these 18 were chosen in such a way that two lists of nine strings, each containing the same number of strings of the same length, could be constructed. The assignment of a string to one list or the other was determined by a table of random numbers."

Using Shannon's measure of channel capacity,* Miller obtains the value 0·552 bits/letter for the redundant strings. For random strings, based on random-numbered tables, he gives the value 2 bits/letter, from $\log_2 4$, four different letters being used in both lists.

The result of decomposition is, of course, the same for both lists, and from the facts that there are 9 strings and 54 letters can be predicted as $\frac{9 \log_2 9}{54} = 0{\cdot}528$ bits/digit. This result is not very different from Miller's for the redundant strings, which was calculated over the 22 strings produced by the generator. Discussing the well-formed strings, Miller says:

"... the redundancy is $1 - (0{\cdot}552/2) = 0{\cdot}724$, or 72·4 per cent. This can be interpreted to mean that the well-formed strings are 3·62 times as long as they would need to be to encode the same amount of information."

The random list (as will be shown later) reduces to a "core" of 20 information-carrying digits, while the redundant list reduces to only 32. This raises the question of the relationship between internal and external redundancy. Miller notes that "in this particular generator, the basic strings themselves seem to group together." And further: "The subjects seemed to distinguish only two basic types of strings." This reminds one of the readily noticeable features of the "principled" selections, and suggests that the finite-state generator may be regarded as a principle of a less simple kind than in those selections.

Samples and Sources

Having described in some detail two types of redundant material, we are now in a position to develop a more general argument. The aim of the psychophysical methods is, of course, to bring into the experimental situation material which can be specified by a non-psychological criterion, preferably quantitative. To use a line "3 in. long" gives complete specification, but to use a line "3 times as long" means nothing unless the comparison line is stated, and unless the ratio has a meaning within the experiment. This is also true of redundancy as a percentage measure between a sample and a source.

Fitts's statement that his constrained patterns are 43 per cent. redundant refers to all "histogram" patterns, of which there are 256. But this is not, in itself, an absolute or an ultimate figure. There is also the expansion of "all patterns made by filling 1, 2, 3, or 4 cells in each column" without constraint to the connected histogram shapes. This gives 30,625 patterns, equivalent to 15·63 bits, a source from which all the Fitts patterns could be selected, but relative to which they have different redundancy measures, the constrained selection being 70·7 per cent., the histogram selection 48·6 per cent.

If two values of redundancy are available, the choice of value requires experimental justification. This, it is suggested, must be to the effect that the chosen source has not merely been contemplated by the experimenter in the process of preparing material, but that in some definite sense it enters into the experimental situation. It will be argued that the "theoretical" values used by Fitts and Miller, both for "redundant" and also "random" material, fail to meet this requirement. Other values will be put forward, for which it will be claimed that they are more relevant values for the material as it enters into experiment.

* In the form $\lim_{\lambda \to \infty} \frac{\log_2 F(\lambda)}{\lambda}$ for strings of length λ or less.

We now ask, as we have asked about the redundant material, in what sense the random material bears the mark of its origin.

In the discussion of discrete noiseless channels which leads up to the definition of redundancy, Shannon (Shannon and Weaver, 1949) state that "In an ergodic process every sequence produced by the process is the same in statistical properties. Thus the letter frequencies, diagram frequencies, etc., obtained from particular sequences, will, as the lengths of the sequences increase, approach definite limits, independent of the particular sequence." Weaver similarly emphasizes that "any reasonably large sample will tend to be representative of the whole."

Clearly the sequences we have considered are not long. In fact, in Fitts's experiment only eight patterns were used in each session, a smaller selection than those here considered. We have seen that, according to its size, the random selection generates transition probabilities: it seems a reasonable suggestion that these are likely to be utilized in the successive processes of memorization and recall. Here the selection as a whole is present, and, by practical necessity, finite. This follows from the fact that it is remembered. What is suggested is that, in so far as the subject can be supposed to act in accord with a hypothesis, it is best regarded as one directed to a source identical in statistical properties with the presented sample. Or to put it more moderately, the subject's hypothesis is directed to the class of sources compatible with the presented sample. The random source will be in this class, but seems pre-eminently unsuited to be the basis of hypothesis, since it is the one source which can give no direction to response. "He adopted the hypothesis of a random source" might serve as an epitaph—perhaps as an alternative to "stimulus deprivation."

The argument started from Fitts's selections because of the clarity of the differences between them. The application to Miller's strings depends upon a re-consideration of Miller's statement that redundant material is longer than it need be to encode the same amount of information.

Miller's experiments

Miller's strings were presented on cards which were shuffled and shown at the rate of one every five seconds. After each trial the subject wrote down in any order all the strings he could recall. Ten trials were given with the same list, some subjects having the redundant lists first, others the random lists. Scoring was by correct strings, incorrect responses being ignored. The first lists for all subjects showed that redundant strings were recalled at each trial in more than twice the number of random strings. In 10 trials (on average) practically all redundant strings were recalled, but only four random strings.

The second list results varied with the type of list first learnt, and showed curious interference effects. These need not be considered as Miller's chief conclusions are based on the first trials.

In a subordinate experiment, described in a footnote, Miller's strings were used in a search task:

> "Lists of 30 and 60 strings were typed on separate sheets and 10 subjects were asked to search for particular strings. . . . It took about twice as long to search through the redundant strings, and it is in this sense that one can say that the redundant strings have a greater degree of similarity."

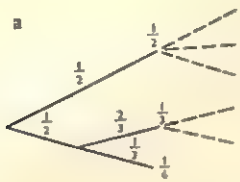
Miller comments on the similarity of this result to that obtained by Fitts with visual patterns, where significantly longer recognition times were obtained for "constrained" than for random patterns. To economize space, we shall take Miller's two experiments as representing recall and recognition situations.

Analysis by Choice Points

The decomposition process, as so far demonstrated, has followed Shannon's principle, but not his method of graphic representation. The diagram of Figure 2*a* precedes his derivation of the $-p \log_2 p$ formula. To it have been added only the final outgoing paths implied by his probability fractions. The decomposition replaces the probabilities with the equivalent uncertainties.

FIGURE 2.

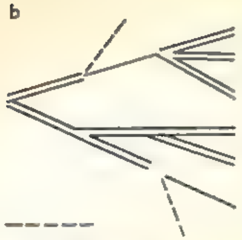
a



11	1·0	1·585		2·585
12	1·0	1·585		2·585
13	1·0	1·585		2·585
211	1·0	0·918	1·0	2·918
212	1·0	0·918	1·0	2·918
22	1·0	0·918		1·918
	6·0	7·509	2·0	15·509

Average uncertainty = 2·585 bits = $\log_2 6$

b



111	1·0	0	1·585
112	1·0	0	1·585
113	1·0	0	1·585
211	1·0	0·918	1·0
212	1·0	0·918	1·0
223	1·0	0·918	0

----- Implied paths

The two forms are of course mathematically equivalent, giving the same analysis when applied to limited sets of items of equal probability such as appear in memorization or recognition situations. The graphic form, however, clearly displays differences of structure between sets which are less obvious in the matrix form. Each choice point (or node) corresponds to a class of similarly occurring symbols in the matrix, and the graph emphasises the number and disposition of these nodes rather than the number of separate symbols.

The subject, whether in a recall or a successive-guessing task is, in effect, tracing paths in a maze towards various terminal points. When he achieves complete recall, he can be said in some sense to know the maze. Suppose that we have six cards bearing the groups of digits of Figure 2*a* and together carrying 14 symbols. A subject who knows the set, and therefore the objective probabilities, is asked to guess at a card from the set. At the start he has to choose between digits 1 and 2, of equal probability. The first digit is uncovered as 2. He now faces an uncertainty of 0·918 bits associated with *any* digit 2 at the choice point: the fact that there are three cards starting with 2 is irrelevant, it is the class symbol 2 on whatever card it happens to appear that marks his choice point and determines his uncertainty. The uncertainty is *at* the node, compounded of the uncertainties on the outgoing paths, which following Shannon we take as the loci of probability. This suggests $\frac{6 \times 2{\cdot}585}{9} =$ 1·72 bits/path as a more relevant figure than $\frac{6 \times 2{\cdot}585}{14} = 1{\cdot}11$ bits/digit; the first figure being sensitive to differences of choice structure, whereas the second is not. One aspect of an efficient coding system is then that each symbol class (or symbol on the probability graph) represents a genuine choice point. In a modification of the

graph (Fig. 2*b*), we can see the structural position of symbols which do not represent choice points. There are now two false nodes which do not correspond to choice points and are meaningless points on the effective paths. The calculated average uncertainty becomes:

$$\frac{6 \times 2{\cdot}585}{11} = 1{\cdot}41 \text{ bits/all paths}$$

$$\text{or} \quad \frac{6 \times 2{\cdot}585}{9} = 1{\cdot}72 \text{ bits/true paths}$$

and this difference is a measure of coding inefficiency in the second form. But it should be noticed that this criterion of coding efficiency can be taken still further. The *most* efficient coding system is always to have a symbol for each item, which gives 2·585 bits/true path *or* bits/symbol. Each symbol is then independent, which is (*cf.* Shannon) the mark of absence of redundancy.

With false nodes and paths we have redundancy in the sense that we are using unnecessary symbols for the task in hand: to discriminate six items. The additional nodes would be justified if there were *at least* sufficient extra items to require them. If we add two such items and make all paths true, including the two new ones, we reach the number of items which that symbol structure is fitted to deal with economically and efficiently. This would cover 8 items, and we define the *empirical redundancy* (E-redundancy) of the original structure relative to the completed structure:

$$\frac{\log_2 8 - \log_2 6}{\log_2 8} 100 = \frac{3 - 2{\cdot}585}{3} 100 = 13{\cdot}8 \text{ per cent.}$$

The structure implied by the false nodes would give:

$$\frac{8 \log_2 8}{13} = \frac{24}{13} = 1{\cdot}84 \text{ bits/paths,}$$

which is still less than the 3 bits/path we should obtain by using the symbols 1 to 8.

E-redundancy offers a measure of a characteristic of redundancy which tends to appear whenever the symbolization is unnecessarily complex for the specific choice task. It is common to all important languages, which have developed as descriptive tools for a wide range of tasks, and are always more complex than is needed for one specific task. The title "empirical" is appropriate both because it cannot be calculated from the general values, and also because the formula has an empirical aspect. (This point will be amplified later.)

A kind of redundancy arises whenever we use a code with fewer symbols than there are distinct items to be denoted. There would be economy of transcription if, for example, we had a separate symbol for each number from 1 to 99, but it would greatly increase and concentrate the discriminatory task. We reduce difficulties of discrimination at the cost of increasing symbol-events, and also spread the risk of total error. The effect is that common elements occur in symbolizations that are distinct in denotation, so that items acquire interdependence and a choice structure. E-redundancy is an attempt to describe that structure for each specific selection.

Analysis of the redundant strings

We shall now apply these principles of analysis to one of Miller's lists, L_1 "a redundant set of well-formed strings." The method of analysis by choice points is less simple in practice than the matrix method because the number of item-termini must be traced ahead, instead of appearing in the next column. At each point the value of $-p \log_2 p$ is

multiplied by the number of possibilities from that point, equivalent to the addition of uncertainties in the matrix. Thus at Step 2:

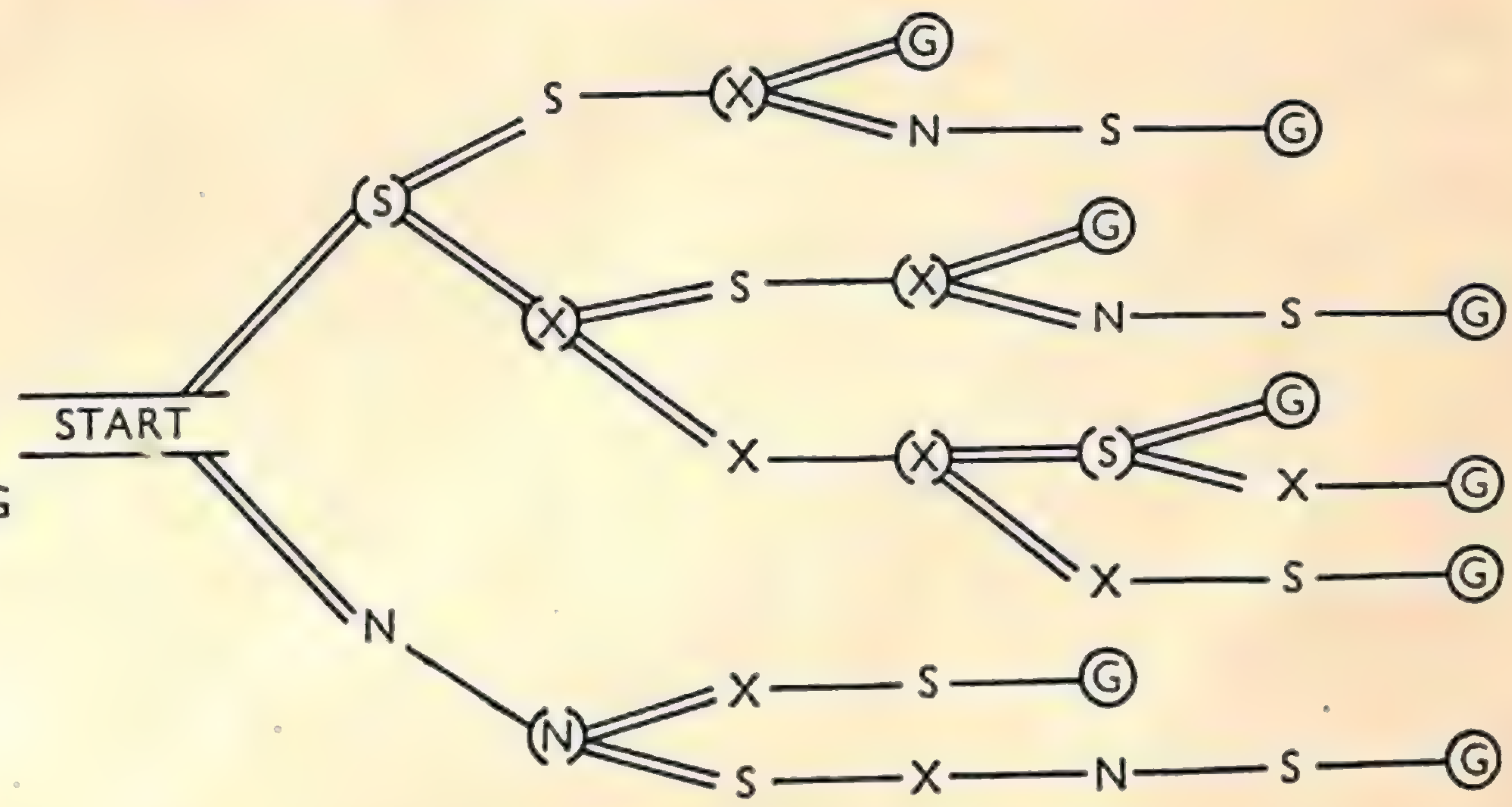
$$\Sigma\text{-}p \log_2 p = 2/7 \log_2 7/2 + 5/7 \log_2 7/5$$

The choice point total is:

$$7(2/7 \log_2 7/2 + 5/7 \log_2 7/5)$$
$$= 2 \log_2 7/2 + 5 \log_2 7/5$$
$$= \Sigma\text{-}n_i \log_2 \frac{n_i}{\Sigma n_i}$$

FIGURE 3

STRING					STEP 0 1 2 3 4 5 6 7
S	SX	G			
S	SX	NSG			
S	X	SXG			
S	X	SX	NSG		
S	X	XX	S	G	
S	X	XX	S	XG	
S	X	XX	XSG		
NN	XSG				
NN	SXNSG				



	0	1	2	3	4	5	6	7
	9x 0·764	7x 0·864	5x 0·971 2x 1·0	2x 1·0	3x 0·918 2x 1·0	2x 1·0	0	0

Step totals 6·876 + 6·048 + 6·855 + 2·0 + 4·754 + 2·0 = 28·533

Giving bits per string = 3·17

= $\log_2 9$

PATHS per step:	2	3	5	6	7	6	4	= 33
No. by false nodes:	0	1	1	4	3	4	4	= 17
TRUE PATHS	2	2	4	2	4	2	0	= 16

Choice nodes are bracketed. Terminal nodes are circled. True paths are double-lined.

Decomposition of Miller's L_1 Redundant List.

The choice graph and decomposition of Miller's L_1 list is given in Figure 3. From the analysis of paths we obtain:—

True paths	16	No. of strings	9
False paths	17	Implied strings	17
	33		26
Implied paths	17	Average uncertainty of implied strings	
Total	50	$= \log_2 26 = 4{\cdot}7$ bits	

$$\text{E-redundancy} = \frac{4{\cdot}7 - 3{\cdot}17}{4{\cdot}7}$$
$$= 32{\cdot}5 \text{ per cent.}$$

Three bits/path figures can be obtained:

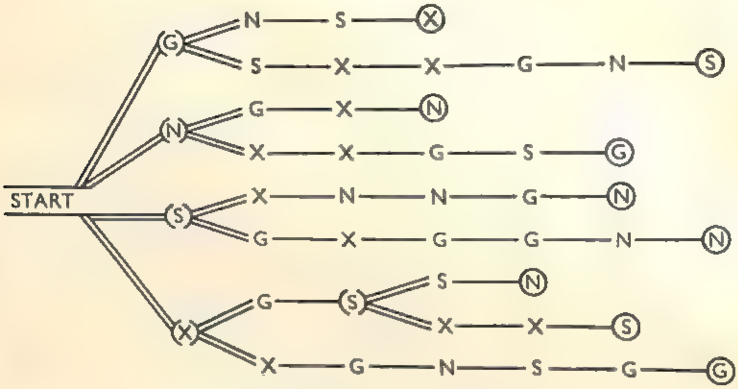
$$\text{Actual} = \frac{9\,(3{\cdot}17)}{33} = 0{\cdot}86 \text{ bits/path (A)}$$

$$\text{Implied} = \frac{26\,(4{\cdot}7)}{50} = 2{\cdot}44 \text{ bits/path (I)}$$

$$\text{Reduced} = \frac{9\,(3{\cdot}17)}{16} = 1{\cdot}78 \text{ bits/path (R)}$$

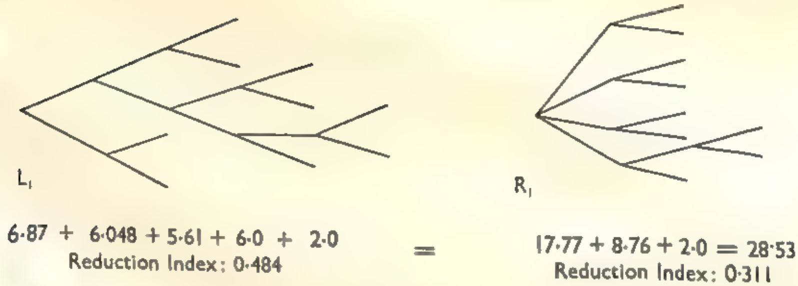
The reduced structure will give full discrimination between the nine strings without E-redundancy. The expression $\frac{\text{True paths}}{\text{All paths}}$, in this case $16/33 = 0{\cdot}484$, gives an index of reduction.

FIGURE 4



Graph of Miller's R_1 Random List.

FIGURE 5



Reduced Graphs.

The graph for R_1, one of Miller's random lists, is given without step-by-step calculation in Figure 4. Results obtained are:

$$\text{E-redundancy} = \frac{5{\cdot}322 - 3{\cdot}17}{5{\cdot}322} = 40{\cdot}4 \text{ per cent.}$$

Bits/path figures:

Actual: 0·634 bits/path (A)
Implied: 2·8 bits/path (I)
Reduced: 2·037 bits/path (R)

By this method of measurement, the redundant L_1 list is less E-redundant than the random R_1 list. The meaning of the new measures can be indicated by comparing the reduced graphs of Figure 5. It will be seen that with exclusion of false paths the random graph reduces to three steps, but the redundant graph reduces only to five. A more efficient four-digit coding could be complete in two steps.

Implications of the Re-analysis

We may now consider the theoretical bearing of the above analysis, and we start from the characteristics of the two graphs as they appear to inspection. In the random list all discrimination points are passed in the first two steps; the remainders of the strings being entirely redundant, without choice points. In the redundant list, choice points are scattered along the strings, and, it might be suggested, provide mnemonic "stepping stones." On this basis we might put forward the qualitative hypotheses that:

(i) For memorization and recall: spreading of choice points is favourable.

(ii) For recognition or discrimination: massing of choice points is favourable.

The second hypothesis is in accord with Miller's and Fitts's results and is also intuitively plausible. In a random list, because it is random, uncertainty tends to be high in the early columns (or indeed in any column taken as starting point), therefore information for discrimination is rapidly available.

The proposed measure of E-redundancy is an empirical one in as much as there is a degree of arbitrariness in using as the reference figure for calculation the *minimum* number of items that would make all choice points genuine. It remains to offer a rational justification of this procedure, in comparison with that used by Miller and Fitts. By their procedure:

(1) A non-random (principled) selection is taken from the ensemble and assigned a redundancy based solely on the number of items obtained, without reference to the principle of selection used.

(2) A further "random" or arbitrary sub-selection is taken as experimental material and assumed to have the same redundancy as the complete non-random selection.

The aim of the present procedure is to state a redundancy for any specific selection taken from the ensemble. Now:

(1) The average uncertainty per item for a selection of n items is, as has been shown, $\log_2 n$, however the items are selected. This figure, therefore, does not give a distinctive measure.

(2) The distribution of uncertainties varies with the specific selection, so that if we could find a general description of the pattern of uncertainties this would give a measure. This we have not so far achieved. But:

(3) As the average uncertainty per item decreases the average uncertainty per digit also decreases, and when the selection is proportionately small this reduction increasingly shows itself in the form of wholly redundant digits (uncertainty $= 0$) with production of false paths.

(4) E-redundancy takes the step of treating the number of false paths as an indicant of the redundancy of the selection. The use of the logarithms of the implied and actual paths is, it may be noted, not based in Information Theory, but is merely a logarithmic scaling which appears suitable, and is open to reconsideration.

(5) Finally, in so far as the "principle" of selection shows itself in the choice graph it is by the patterning of the uncertainties. Now it can be shown (though only at some length) that the measure of E-redundancy offered varies with the factors that tend to produce spreading of the choice points, so that (as in the cases shown) high E-redundancy goes with massing of

choice points, low E-redundancy with spacing of them. It is this that suggests that E-redundancy can be used as an indicator of the form of the choice graph, though its relevance to task situations can only be established experimentally.

Experimental results

In an uncompleted series of experiments, two selections were taken from Miller's 22 redundant strings, such that their E-redundancy varied considerably. It was predicted that those with lower E-redundancy would be more easily recalled.

The conditions of the experiment were as for Miller's first trials, except that eight items were used instead of nine. This was because the same items were also to be used in a recognition apparatus similar to Fitts's, taking only eight items. Fourteen subjects were used in two groups which received the lists in opposite orders.

The average recall for the A list (E-redundancy 30·58 per cent.) was 6·92 items; for the B list (E-redundancy 37·58 per cent), 4·72 items. The individual scores were as follows:

	A list first (seven subjects)							B list first (seven subjects)						
A list:	6	8	8	7	7	6	6	7	8	5	8	7	8	6
B list:	5	7	3	3	7	3	8	4	4	4	5	6	3	3

It will be noted that only one subject gave a result contrary to the tendency of the pooled results. (A later repetition with 16 different subjects has given similar results.)

These results suggest provisionally:

(i) That a predictive distinction can be made between lists which by Miller's measure are of equal redundancy.

(ii) That the form of the choice graph has relevance.

(iii) That E-redundancy may be a useful assessment of this form.

More adequate experiments on both recall and recognition are in progress.

The opposition between Miller's measure of redundancy and E-redundancy suggests that fundamentally different concepts of redundancy are involved. These would seem to correspond to two aspects of redundancy which can be found in Shannon's formulation. In the "noiseless" channel any symbols or message units which over-determine decision are merely useless. They could be omitted or replaced with any neutral symbol. The matrix of uncertainties defines the relative usefulness (or uselessness) of symbols for the purpose of recognition, and it is this average per item which is given by $\log_2 n$.

With noise, however, symbols which are useless in the noiseless channel may acquire uncertainty-reduction values of the obscured signals which they replace. This "useful redundancy" derives from interdependence of symbols or units, and, it is suggested, from the building in of a "principle" which is not in itself defined by information measures.

E-redundancy then, makes no claim to define "principle," but only to serve as a measure of symbol interdependence or structure of choice in a finite selection of meaningless material. In the writer's view the effect of principle in memorization and recognition tasks remains to be explored by other means. The immediate conclusion is that finer specification of task material is both possible and essential. A broader conclusion, based both upon this enquiry, and upon the discussion of redundancy in visual forms by Attneave (1954), is that the word "redundancy" is carrying too many and too varied burdens, some of which touch on problems of meaning without adequately analysing them.

A last comment may be made on Miller's redundant list. As it was presented to the subjects it was meaningless and therefore only discrimination criteria can be

applied to it. But it can be given meaning if we regard the strings as, e.g. prescribed racing-car routes on the finite-state map. In this interpretation no symbol is redundant for a driver, provided that he does not already know the constraints on entry to certain roads. If he knows these, or if the roads are labelled "ONE-WAY: NO ENTRY," a coding nearer to discrimination coding would suffice. Thus, finally, redundancy in a string of symbols can only be defined relative to *a priori* information and the message value for the recipient.

Error-combating redundancy is of course fully discussed by Shannon, and clearly distinguished from other redundancy. A rough criterion is that the decoded message retains language redundancies, but the transmissive redundancy is removed to obtain the message. Shannon, however, does not cover the case (common in experimental situations) of the human receiver who does not know what he is trying to receive. This remains a strictly psychological problem to which the key is probably still to be found in Bartlett's (1923) "effort after meaning."

The author is indebted to Dr. J. Szafran for much helpful discussion and criticism.

References

ATTNEAVE, F. (1954). Some informational aspects of visual perception. *J. exp. Psychol.*, **61**, 183–94.

BARTLETT, F. C. (1923). *Remembering*. London.

FITTS, P. M., WEINSTEIN, M., RAPPAPORT, M., ANDERSON, N., and LEONARD, J. A (1956). Stimulus correlates of visual pattern recognition: A probability approach. *J. exp. Psychol.*, **56**, 485–91.

MILLER, G. A. (1951). *Language and Communication*. New York.

MILLER, G. A. (1958). Free recall of redundant strings of letters. *J. exp. Psychol.*, **56**, 485–91.

QUASTLER, H. (Ed.) (1955). *Informatio ' Theory in Psychology*. Illinois.

SHANNON, C. E., and WEAVER, W. (194 . *The Mathematical Theory of Communication*. Illinois.

PERCEPTION OF SEQUENCE IN AUDITORY EVENTS

BY

PETER LADEFOGED and DONALD E. BROADBENT

From the Phonetics Department, University of Edinburgh, and Applied Psychology Research Unit, Cambridge

A series of tape-recorded sentences were presented to various groups of listeners, totalling 164. During each sentence an extraneous sound was present on the recording, and the listener had to indicate the exact point in the sentence at which this sound occurred. It was found that errors were made which were large compared with the duration of a single speech sound; which suggests that the listener does not deal with each sound separately but rather with a group of sounds. Errors were reduced if the sentence consisted of a series of digits rather than an ordinary text, or if the listeners were trained in phonetics. Prior knowledge of the content of the sentence did not affect accuracy. The direction of error was usually to refer the extra sound to an early point, but it is affected by the relative position of the extra sound in the sentence. These results can be regarded as an extension, to the case where all stimuli are presented to the same sense, of classic results on prior entry.

Introduction

This paper presents data which suggest that the central processes involved in understanding spoken language may have certain similarities with those involved in understanding written language. These similarities are often overlooked, because listening to speech is usually supposed to be a process in which the different speech sounds are recognized one after another. Each recognition is of course regarded as being influenced by past ones, so that a sound is more readily perceived when it is in an appropriate context. But the units on which the process is supposed to operate are single speech sounds. In reading however, it is well known that the movements of the eye do not involve fixating on each letter in turn, but instead consist of a number of rapid saccadic movements over several letters, with comparatively long pauses in between each jump. However, this method of scanning the visual data is not due to limitations of motor skill (since movements much smaller than a normal saccade can be made) nor to lack of sensitivity of the retina (which is adequately stimulated in far less than a normal fixation pause). It is a central rather than a peripheral process which is responsible for each section of the visual material being sampled for an interval of time rather similar to that necessary for the central process in a simple reaction time experiment. This being so, it would not be surprising if auditory data were sampled in a similar way. That is, if several successive sounds were handled together as a group by the mechanism responsible for identifying them. Yet if this were so the physical order of arrival of sounds might not determine their perceived order. As one sound would not be recognized and disposed of before the next arrived, all the items in each sample would be effectively simultaneous and any of them might be perceived as the first.

The experiments to be reported are not conclusive but they tend to substantiate this hypothesis. They are all based on the fact that observers find it very difficult to locate the time of occurrence of a short sound superimposed on a recording of continuous speech. The original observation made by one of the authors (P.L.), is that although he is a trained phonetician accustomed to listening to speech analytically, he was unable to note exactly when a switching click occurred in a recorded utterance which was being played back over and over again on a loop of tape.

CLICK EXPERIMENTS

Experiment I

Two groups of 12 Naval Ratings each heard a recording of a story consisting of the following eight sentences superimposed on each of which, at the point marked by a vertical line, was a click with an intensity approximately equal to the intensity of the vowel sound with the maximum stress in the sentence.

(1) There was once a young r/at named Arthur.

(2) Who could never m/ake up his mind.

(3) Whenever his friends invite/d him to go out with them.

(4) He couldn't think whether he ought t/o go or stay at home.

(5) And on the day when all the o/ther rats moved to a new house.

(6) Where they thought they'd be able to f/ind more food.

(7) Arthur couldn't bring himself t/o a decision of any kind.

(8) So he stayed behind until eve/ntually he starved.

The subjects heard a specimen click and in each sentence were told to note the words on which or between which the superimposed sound occurred. The first group did not know the story at all, and had to write down the appropriate word or words. The second group knew the story, had it written in front of them, and had each sentence read out before being played. This group noted the appropriate word or words on the text which had been provided.

Experiment II

A third group of 10 Naval Ratings heard the same recording, with instructions to mark the position of the click on a text provided. They also heard a recording of the following 10 numbers.

(1)	7	1	9	2/	6	5	10	3	8	4
(2)	3	10	5	9	2	7	1/	8	4	6
(3)	9/	1	4	6	5	10	3	8	7	2
(4)	2	6	3	10	4	8	9	1	7/	5
(5)	1	4/	9	5	3	6	10	2	8	7
(6)	9	8	1	7	10	3/	5	4	6	2
(7)	6	8	7	1	5/	4	2	10	9	3
(8)	10	7	2	3	1	9	6	5	4	8/
(9)	8	2	6/	4	9	1	7	5	3	10
(10)	4	5	8	6	3	9	2	10/	7	1

Each number was said at an even rate with an equal stress on each digit. A click similar to those in Experiment I was superimposed on each number at the point shown by the vertical line. Subjects were asked to mark on answer sheets which were provided the digits on which or between which the superimposed sound occurred. Half the subjects heard the number recording before the story, and half received the opposite order.

General procedure and scoring

The recordings were made on Ferrograph tape recorders (frequency response 50–10,000 c.p.s. $\pm$ 2 db.) and played back on similar machines in an ordinary reverberant room. Testing was carried out in groups. Spectrograms were made to locate the superimposed items accurately. The answers of each subject were scored in the following way. Each word and each space between words was called a position. If the subject had indicated the position where the superimposed item actually occurred, his score for that sentence was zero: if he put his mark in the next earlier position, plus one was counted. If he marked the position before that, the score was plus two. Marks placed after the correct position were scored minus in a similar way. For example, in the first sentence of Experiment I, a score of plus four would mean that the word "a" had been marked, while a score of minus three would mean that the mark was between "named" and "Arthur."

Results

Experiment I. The average score per individual per sentence for the last six sentences of the story was 1·83 for Group 1 (not knowing the story in advance); and 2·3

for Group 2 (thoroughly conversant with each sentence of the story). The difference between the groups is completely insignificant statistically: but each score is highly significantly different from zero. (All subjects, except one in Group 1, had a positive score.) The click tends in fact to be heard on the average as occurring during the word before the one which it accompanies. This inaccuracy is all the more striking when we recall that each word consists of several sounds in an order which, one might think, must be perceived if the word is to be recognized. The absence of a significant difference between groups means that we have no evidence for any effect of prior knowledge of the text upon the ability to locate a superimposed item.

Experiment II. If we take the average error per sentence, ignoring sign, per subject in the third group we find that it is 2·0 for the story and only 0·7 for the numbers. This difference is highly significant, as every subject did better on the numbers. Thus there seems to be something about an ordinary sentence which makes it harder to judge the location of an item superimposed on it.

In addition, the fact that the different numbers provide examples of clicks in all possible positions, from beginning to end of the stream of sounds, allows us to check that the bias towards an early location for the click is not due to the particular locations used in the story. In the numbers, all positions of the click gave an early subjective location, with the one exception of the click in the fourth position.

We cannot therefore explain the bias towards an early location as due, for example, to a tendency to place the click nearer the middle of the range of possible positions. But in these particular subjects the earliness was more marked when the click came later: the tau correlation between the actual position of click and the size of the error in estimating the position is 0·58 ($p < 0·05$). As this correlation is due largely to four subjects we cannot say confidently that it would appear in a fresh sample of subjects.

SPEECH SOUND EXPERIMENT

Experiment III

A possible explanation of the results given so far might be that the click used was a short and mechanical sound and not one which would in normal experience form a meaningful part of a speech sequence. Listeners might be better able to locate a sound which was exactly the same as one of the normal sounds of speech, and which could not in the first instance (i.e. before any of the sounds were identified as belonging to meaningful groups) be regarded as noise or an isolated sound having no further need of processing. The click experiments also made no attempt to study the effect of putting the extra sound in various relations to the sense group in the speech. The following experiment covers these points.

A total of 130 students heard the following set of five sentences, superimposed on each of which, in the place indicated by the vertical lines, was an [s] sound.

(Jean) He *thought* it was *Jean* w$\left|\begin{smallmatrix}s\\h\end{smallmatrix}\right|$o had a *cold* and was in *bed*.

(Jane) We *thought* it was *Jane* who could $\left|\begin{smallmatrix}s\\ \ \end{smallmatrix}\right|$ be *brave* and in the *team*.

(Jim) He *hoped* it was *Jim* who could be *k*$\left|\begin{smallmatrix}s\\e\end{smallmatrix}\right|$*pt* when he was *old*.

(Jack) It's *lucky* that *Jack* was in the *boat* that $\left|\begin{smallmatrix}s\\ \ \end{smallmatrix}\right|$ had a *leak*.

(John) They *think* it was *John* that was the *boy* that had $\left|\begin{smallmatrix}s\\a\end{smallmatrix}\right|$ *top*.

The prosodic features such as stress and intonation were similar in all these sentences, each being said with a heavy stress on the italicized words. The rate of delivery was that of a normal conversational style. The superimposed item was an [s] sound with a

duration of 40 millisec., and comparable in intensity with the [s] in *It's* in the *Jack* sentence. (This is the only [s] in these sentences; the other sounds which are written as *s* are phonetically [z].) The recordings were made on Ferrograph tape recorders, with the aid of a Vortexion mixer unit. The overall frequency response was substantially flat over the whole speech range (± 3 db. from 50 to 9,000 c.p.s.). Sound spectrograms were made of the sentences, so as to check that each superimposed item was in exactly the intended position. Where an item was to occur on top of a word, care was taken to ensure that part of the vowel in the word occurred on either side of the [s]: and when it was to occur

TABLE I

ANALYSIS OF VARIANCE FOR THE FIVE GROUPS RECEIVING SENTENCES IN DIFFERENT ORDERS IN EXPERIMENT III

Sources of variance	*d.f.*	*Sum of squares*	*Variance*	*F*
Between Groups	4	69	17·25	<1
Within Groups	55	3429	62·34	
Total	59	3498		

TABLE II

ANALYSIS OF VARIANCE FOR THE GROUP OF TRAINED LISTENERS AND THE 60 UNTRAINED SUBJECTS IN EXPERIMENT III

Sources of variance	*d.f.*	*Sum of squares*	*Variance*	*F*
Between Groups	1	336	336	5·7
Within Groups	67	3952	58·98	
Total	68	4288		

TABLE III

THE MEAN JUDGEMENTS OF THE GROUP OF TRAINED LISTENERS AND OF THE 60 UNTRAINED SUBJECTS IN EXPERIMENT III

No. of subjects	*Composition of group*	*Type of sentence*					*Group totals*
		Jean	*Jane*	*Jim*	*Jack*	*John*	
9	2nd year speech therapists	−3	+2	+1	0	0	0
60	Subjects in Table I	−1	+3	+1	+1	+2	+6

between words it was arranged to be during a stop closure. This closure belongs phonetically to one word or another; in ordinary speech there are seldom gaps marking words boundaries. But a stop closure such as that of the [t] in the phrase *that had* results in a period of silence after the last audible part of one word and before the first audible part of the next word: an item superimposed during such a closure may thus be said to occur between words.

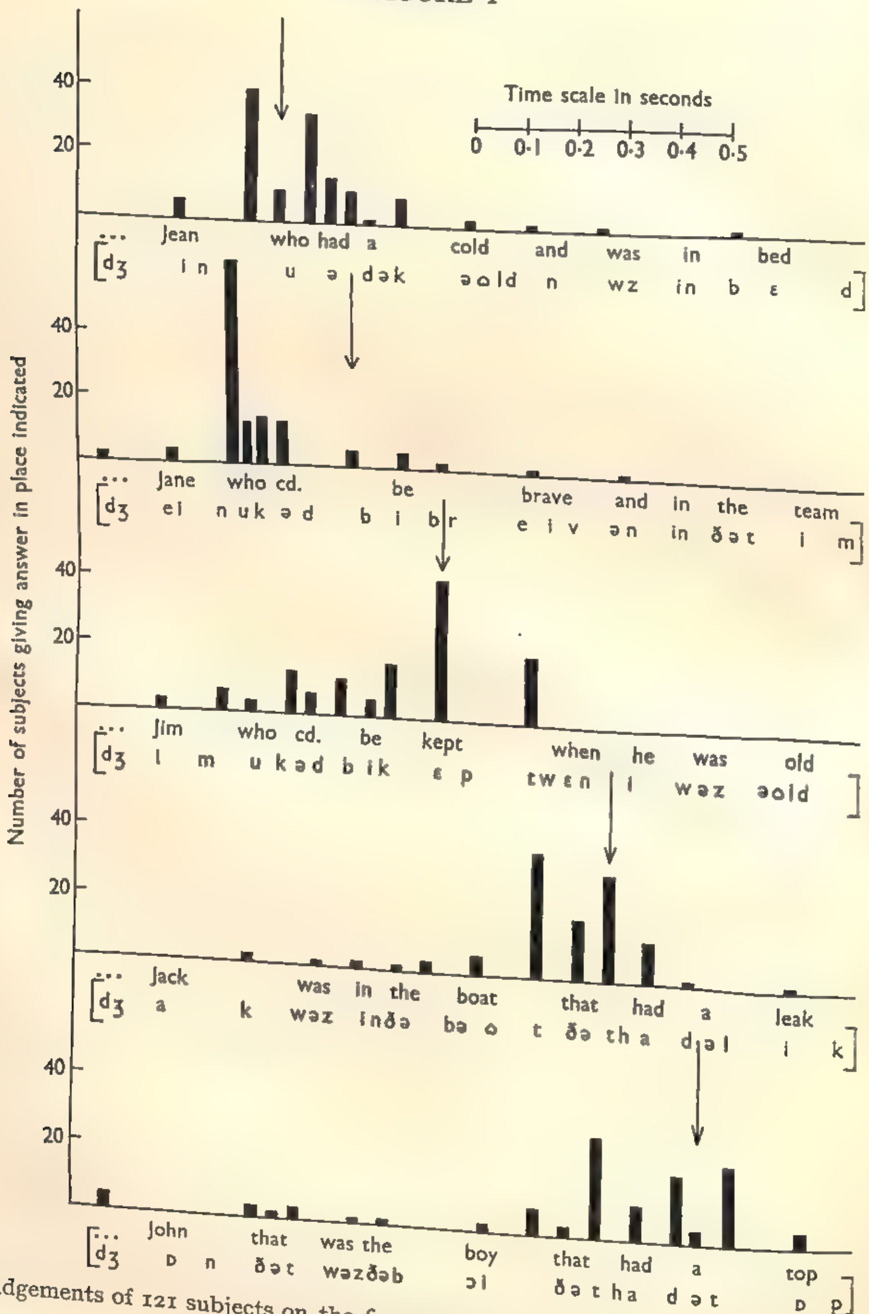
All the subjects in this experiment were given a text of the five sentences which they read while first listening to the sentences without any superimposed items. They were then given written and duplicate oral instructions explaining that they were going to hear the same recording again, but this time there would be an [s] sound superimposed on each sentence; and they were to mark on the provided texts the words on which or the

words between which these sounds occurred. They were instructed not to read the text at the same time as they listened to the recordings.

The 130 subjects were considered in different ways.

(*a*) 60 of them were divided into five groups, each of which heard the five sentences in a different order forming a Latin Square. (*b*) Nine of the remaining subjects were students who had been taking classes in Phonetics for over a year. (None of the other subjects in

FIGURE 1



The judgements of 121 subjects on the five sentences. The actual position of the extra sound is indicated by an arrow

any of the experiments had been trained to listen to speech analytically.) (*c*) The remaining 61 subjects who were not used for statistical analysis all heard the sentences in the order *John, Jane, Jean, Jack, Jim*.

Results

(*a*) *The five groups who heard the sentences in different orders.* For each subject four scores were obtained by taking away his score on the first sentence from that on the second, third, fourth, and fifth.

A multivariate test of the equality in different groups of these four scores was then performed. (This test, which is based on Hotelling's T^2, was performed by Dr. M. Stone, of the Applied Psychology Research Unit). The test showed that the four scores were different in the five groups, $p < 0{\cdot}01$.

This must mean either that the different sentences are different in the type of error they produce, or else that the different sentences produce different amounts of practice. That is, either a sentence such as *Jim* produces a different error from some other sentence, or else judging the *Jim* sentence after the *Jean* sentence is different from judging the *Jim* sentence after the *Jane* sentence. The latter possibility was tested by taking the total score for each subject, counting all five sentences. Analysis of variance was then performed on these total scores, as shown in Table I. The groups were not significantly different. Thus the total score on all five sentences is not different if they are heard in different orders. Therefore the previously proved difference in the four scores in the five groups is likely to be due to the five sentences producing different types of error.

(*b*) *The group of trained listeners.* The total score for each subject in this group was determined, counting all sentences. The difference between the group and the 60 other subjects already described was then tested as shown in Table II. This group is therefore different from the others ($p < 0{\cdot}05$). The mean judgements of this group and the other 60 subjects are shown in Table III. It seems that this group is different in that it is often slightly more accurate.

(*c*) *The remaining* 61 *subjects.* The responses of these subjects taken together with those of the 60 subjects already discussed are shown in Figure 1. The words are displayed (in orthography and in a phonetic transcription) in accord with their duration in the utterance.

Discussion

All these experiments show that subjects have a great deal of difficulty in identifying the time of occurrence of some auditory events relative to others; and that they often consider the item which they are trying to pick out to come earlier than is actually the case. This situation is suggestively reminiscent of the early experiments on "prior entry" in which, e.g. subjects are told to watch a pointer moving around a drum, and to listen for the sound of a bell. In these circumstances subjects usually judge the bell to have sounded at a point which was passed by the pointer somewhat earlier than the actual point at which the bell was struck. This effect is not solely due to the shorter sensory conduction time of the ear since if subjects are instructed differently, and do not expect to hear the bell, they will, if asked afterwards, consider that it was struck at a point somewhat later in the pointer's movement than was actually the case. Titchener (1909) formulates the resulting "law of prior entry" as follows: "The stimulus for which we are predisposed requires less time than a like stimulus, for which we are unprepared, to produce its full conscious effect."

All the early work on this "law" involved stimuli affecting different sense organs, and it is of course easy to think of information in one sensory channel as being

delayed relative to that in some quite different channel. But if the phenomenon of prior entry applies even to items all arriving by one sense, we must conclude that these items do not pass along the sensory paths in rigid succession and to that extent the concept of sensory sampling is justified. It is therefore worth assessing the extent to which the results described in this paper can be considered as being similar to the older experimental data on prior entry. In particular, can the shifts in location of the extra item, produced by changes in the speech material, be regarded as similar to the changes produced by "predisposition" in the classical experiments?

Some of the results can be reconciled with the earlier ones. It is a plausible hypothesis that the extent of a subject's predisposition to respond to a class of stimuli is related to the amount of attention which he is giving to that class of stimuli; and this will depend on his estimate of the probability that some member of the class is about to occur which will convey information essential for a response. Subjects will be paying more attention to a group of stimuli when there is a high probability that an information bearing stimulus is about to occur. Thus reaction time is shorter when a stimulus is presented at a time when it is highly probable (Mowrer, 1940); and conversely reaction to visual stimuli is slower the more information is conveyed by simultaneous auditory signals. Now in Experiment II we compared meaningful sentences and sequences of digits as the speech material. We are handicapped by a lack of knowledge about the amount of information conveyed by ordinary sentences such as those in the story. But there are certainly marked sequential dependencies in English sentences which are lacking in the digit series; and the sentences are easier to remember than the numbers. It is therefore probable that the sentences do not carry information at as great a rate as the series of digits, in which lexical information can be conveyed at the rate of 3·1 bits per word. Consequently, when decoding the information in the sentences in the story, the subject does not pay as much attention to the words in the utterance as he does when decoding the information in an utterance consisting of a series of digits. In the latter case his predisposition, which is normally towards the superimposed item, will be reduced: and therefore it is reasonable that he will be more accurate when the speech material consists of a series of digits.

The other finding in Experiment II was that the subjects studied showed a larger error when the click was later in the series of digits. This may plausibly be explained by the rise in the probability of the click as each series continued, since there was bound to be a click at some point, and if it had not happened in the early stages it was certain to come in the later ones. Thus the listener would be more predisposed towards the click when it came late. An alternative theory might be that this result was due to the central tendency of judgement, common in psychological situations. This would not exclude the operation of prior entry but be additional to it.

Similarly some aspects of Experiment III are in agreement with a modified prior entry theory. Each of the sentences in Experiment III was constructed so that it had three sense groups, each ending with a stressed word. In order to decode the information conveyed by such sentences a great deal of attention will be needed both at the beginning of each sense group, and at the moment of occurrence of stressed words, since at these times the message is relatively less predictable. The *Jean* sentence is the only one in which the extra item occurs on the first word in a sense group; and it is also the only one in which a considerable number of subjects considered the item to have come later than it actually did; so perhaps at this moment in the sentence subjects were paying more attention to the words than at other moments.

When the superimposed item comes on a stressed word (i.e. in the *Jim* sentence) presumably many subjects are paying about equal attention to the words and to the

extra sound for which they are listening, since the *Jim* sentence was the one which many (42) subjects got right. It is interesting to note in connection with this sentence that there is a rule of English prosody: *stresses tend* to re*cur* at *reg*ular *in*tervals of *time*. Thus if the last half of the previous sentence is read aloud in a normal conversational style, while tapping a pencil on the table at the moment of saying each stressed (italicized) syllable, it will be found that, almost irrespective of the number of unstressed syllables between each of the stressed ones, the taps recur at regular intervals. Consequently, in many English sentences, after hearing the first two or three stresses the listener can approximately predict the moment of occurrence of the subsequent stresses; so the times when the maximum information will occur, and hence the times when the maximum attention will be needed, are to some extent known in advance.

It should be noted that the results of Experiments II and III argue strongly against a possible alternative to the "prior entry" explanation. This alternative view is that speech takes longer to be perceived than an isolated sound does, because it involves more complex processes. If this were so we would expect the superimposed sound to be regarded as coming even earlier when the simultaneous speech is less redundant, since the perception of the speech should be even further delayed (Hick, 1952; Hyman, 1953). But on the contrary the location of the superimposed sound appears later and so more accurate when the speech is less constrained. Thus the error is not likely to be caused by delays due to the complex character of speech.

The most serious difficulty, in relating these results to earlier prior entry effects, lies in Experiment I. The finding in that experiment was that knowledge of the text of the story in advance did not make any significant difference to the degree of difficulty of identifying the relative time of occurrence of the superimposed item. If knowledge of the text reduced the amount of information decoded when listening to the story, we would have expected one group to be more accurate than the other. This difficulty can perhaps be met by recalling that there is no reliable estimate of the amount of information conveyed by a spoken message as compared with the amount of information conveyed by the same words in written form or spoken by another voice. It is quite probable that in the conditions of these experiments prior knowledge of the lexical structure of the text has little bearing on the amount of information conveyed by the utterance as an auditory event: knowledge of the words that a speaker is going to use is only of slight assistance in predicting the actual noises—including the rate of speech, the stress pattern, the intonation and the voice quality that are going to occur. So it is perfectly possible that the subject has to attend to the sounds of the words to an equal extent, whether he knows the text or not.

Another and more minor difficulty is that the differences between the *Jane, Jack* and *John* sentence of Experiment III cannot readily be accounted for on a modified prior entry view. However, the interesting fact, which is applicable to all the sentences in all the experiments, is thoroughly established: subjects often have great difficulty in assessing the relative time of occurrence of different auditory events, even when they have been specifically asked, before hearing the events, to try to do so. This difficulty is present both when the auditory events are of different kinds (i.e. speech sounds and a click) and when the auditory events are all of the same kind (i.e. all speech sounds).

We may conclude from this that the process of decoding information when listening to speech may involve operating on units which are somewhat larger than the duration of a single speech sound. Admittedly it is not possible to say much more about the units, for there is no definite maximum error in the location of superimposed items. The auditory samples cannot therefore begin and end at the same place for all listeners to a particular set of sounds. Thus the auditory psychologist

is in a similar position to that of his visual colleague. Carmichael and Dearborn (1948) says: "From the study of a great many eye-movement records the conclusion is drawn that the eye seems to have no particular preference for fixating any special part of a word or type of letter or length of word." It appears that reading and listening are similar activities in that they both normally involve shifts of attention in which there is a large random component. Despite this randomness, it is certain that reading is not a letter-by-letter process: and equally listening does not deal successively with each isolated speech sound.

References

Carmichael, L., and Dearborn, W. F. (1948). *Reading and Visual Fatigue.* London.

Hick, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11–26.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, **45**, 188–96.

Mowrer, O. H. (1940). Preparatory set—some methods of measurement. *Psychol. Rev. Monogr.* No. 52.

Titchener, E. B. (1909). *Lectures on The Experimental Psychology of the Thought-Processes.* New York.

A CONTRIBUTION TO THE STUDY OF PHENOMENAL CAUSATION*

BY

D. G. BOYLE

From the Department of Psychology, University of Hull

In his studies of the perception of causality, Michotte describes a number of "impressions," concomitant with precisely defined experimental conditions, of which "Launching"† and "Triggering"† are two of the most important. Two major factors influencing the nature of the impressions are the relative speeds of the two rectangles representing objects participating in dynamic inter-relationships, and the length of path of the more passive of the two objects. This investigation was an attempt to discover the quantitative relationships obtaining between these two independent variables and the dependent variable of the subjects' responses to the experimental conditions.

Introduction

By means of an experimental device to be described in the following section, Michotte arranges for an observer to view two rectangles, of sides 5 by 8 mm., of which one (black) appears to move across the field of view from left to right, and make contact with the other (red), which up till then has been stationary in the middle of the field. After being struck, the red object moves away in the opposite direction, either at a lower speed than the black object approached it or at a higher speed.

The colours of the rectangles are irrelevant to the impression. Whatever colours are used, the object that first approaches from the left is known as object A, and is regarded as the more active of the two. The object that moves away after being struck is known as object B, and is regarded as the more passive.

Briefly, if the speed at which A approaches B is greater than the speed at which B withdraws after impact, the impression is that A has "launched" B into motion. This impression Michotte calls "Launching." If, on the other hand, B withdraws after impact at a speed greater than that at which A approached it, the impression is that A has "triggered" B into motion. This impression Michotte calls "Triggering." The combinations of speeds, and the resulting modifications of impressions, are very great, but we have described the simplest cases as they affect the experiments to be reported.

Generally speaking, as the length of B's path increases, the impression is more likely to be Triggering than Launching, even if the ratio of speeds A : B is greater than unity (i.e. when A approaches B more quickly than B subsequently withdraws.) This case will be known as a "descending ratio" of speeds. When the length of B's path is fairly short, the impression is more likely to be Launching than Triggering, even when the ratio of speed A : B is less than unity (i.e. when B withdraws after impact more quickly than A approached it.) This case will be known as an "ascending ratio" of speeds.

These observations refer to preliminary experimentation conducted at Louvain by Professor Michotte and the present author during the session 1956–57. They are related to the concept of the "radius of action," that is, the distance that B may

* This work was conducted at Louvain, under the direction of Professor A. Michotte, during the session 1956–57, and was first reported in the author's Ph.D. thesis presented to Hull University in July, 1958.

† The terms used in this paper are those adopted by Michotte's translators, T. R. Miles and Mrs. Miles, of the University College of North Wales, Bangor.

travel while still appearing to be under the influence of A. Beyond a certain point, B's movement becomes phenomenally autonomous. This point cannot be discussed at any length here; the reader is referred to Michotte (1954) and to the experimental studies of Yela (1954) for further details. The concept must be mentioned, however, as the radius of action put an upper limit on the length of path that we were able to use in the present experiments.

Altogether, about 80 subjects took part in the present set of experiments. Of these, 40 were selected for the major experiment. They were asked to respond (in the method to be described in the following paragraph) as the length of B's path was systematically increased or decreased under the conditions of either an ascending ratio of speeds or a descending ratio (four sets of conditions altogether.)

The nature of the experimental response differed in the present major experiment from the nature of the responses in Michotte's experiments. Michotte placed no restrictions on the nature of his subjects' responses, telling them to describe what they saw "as if they were telling a friend something that they had seen happen in the street." For the present experiment, the subjects were asked to compare the experimental stimulus with two standard stimuli, one of which unequivocally gave them the impression of Launching, whilst the other unequivocally gave the impression of Triggering.

The experimental hypothesis was that, in the case of an ascending ratio of speeds, lengthening B's path would tend to reinforce the forces of organization, leading to a greater tendency to report Triggering at greater lengths of trajectory. In the case of a descending ratio of speeds, increasing the length of trajectory would lead to a conflict of the forces of organization, leading to ambiguity of response at greater trajectory lengths.

Apparatus

A full description of the apparatus is given by Michotte (1954), and summarised by Vernon (1952); a detailed description would take a considerable amount of space, and would require diagrams. Since such a description would be out of place in this paper, and since details can be obtained easily from Vernon (1952), we shall describe the apparatus only briefly.

The basic apparatus consists of a circle of thick white card, 48 cm. in diameter, on which are drawn two bands 8 mm. in width; one of these is black, the other red. (In practice, two discs are usually employed, each cut along a radius and one set of bands drawn on each. These can be fitted together, making it possible to employ different combinations of speeds without the necessity of drawing a new disc each time.)

The discs are then mounted on plates of the same diameter, revolving at constant speed in the vertical plane behind screens. In each screen a slit is cut, which is 5 mm. high and about 150 mm. long. The observer, viewing the apparatus from a distance of 1·5 metres, sees each band through the slit as a rectangle of side 5 by 8 mm.

If the bands are drawn along arcs of constant radius, the observer will see stationary "objects," even when the discs are revolving. If, however, the arcs are of decreasing radius, then, when the discs revolve, the objects will appear to approach the centre of the slit; if the arcs are of increasing radius, the objects will appear to move away from the centre. In this way it is possible to create the appearance of object A approaching object B, which is stationary at the centre of the slit, then hitting B, after which B moves away, A then remaining stationary.

A series of preliminary experiments was conducted to discover the speed ratios and lengths of path that would yield the clearest-cut results. As well as the stimulus conditions chosen on the results of these investigations, two other sets of stimuli were necessary.

The first of these has already been referred to, viz. the stimuli that gave unequivocal impressions of Launching and Triggering to the subjects taking part in the major experiment. We shall refer to these as "Prototype stimuli." The other set was necessary because most of the subjects employed were naïve to this type of experiment, and in many cases to psychological experimentation in general. It was therefore necessary to explain

to them the *type of response* that was required in this series of experiments. Consequently preliminary training was necessary to ensure that subjects were attaching similar meanings to the permitted experimental responses.

The stimulus conditions were as follows.

Prototype stimuli

Speed of A (cm./sec.)	4·5	Speed of B (cm./sec.)	30·0
" "	30·0	" "	3·0

Training stimuli

Speed of A (cm./sec.)	18·0	Speed of B (cm./sec.)	4·5
" "	9·0	" "	12·0

Experimental stimuli

Speed of A (cm./sec.)	45·0	Speed of B (cm./sec.)	30·0
" "	13·5	" "	30·0
" "	54·0	" "	22·5
" "	9·0	" "	15·0
" "	30·0	" "	15·0
" "	6·0	" "	15·0

The apparatus was arranged with a movable flap fitted to the slit in the screen, so that the length of B's path could be increased or decreased by the experimenter. A scale was drawn on the screen so that the path could be altered by fixed steps of 5 mm. This scale was drawn so faintly as to be invisible from the subjects' viewing position, 1·5 metres in front of the apparatus, especially when this was frontally illuminated.

A number of revolving plates were employed, so that the discs did not have to be changed with each change of stimulus condition. The stimuli not being presented were covered until required by movable screens. The subject moved from one stimulus condition to the next, viewing each frontally from the same distance. All discs were clearly illuminated by frontal illumination.

Procedure

(1) *Training and pilot experiments*

All naïve subjects were presented first with the prototype stimuli and asked to describe their impressions. No specific instructions were given as to the aspects of the situations on which to concentrate. Subjects who described the prototype stimuli in the way expected from Michotte's reports (i.e. those who described the stimulus with A : B = 30·0 : 3·0 in terms appropriate to Launching, and the stimulus with A : B = 4·5 : 30·0 in terms appropriate to Triggering) were selected for training. It was found that about 50 per cent of the subjects had to be rejected on the basis of this preliminary screening test, either because their replies were completely ambiguous, or because they described the stimuli in ways opposite to that expected.

The selected subjects were then shown one of the training stimuli with object B's path reduced to 5 mm. (the path was measured from the edge of B nearer the slit, so subjects saw B as a red square of 5 mm. side, with no free space behind it.) They were asked to say which of the two prototype stimuli the presented stimulus resembled more. If they could not make up their minds they were allowed to say "uncertain," "in between" and so on.

After subjects had responded, the length of B's path was increased by the experimenter to 15 mm. and subjects were asked to make a further judgement of "more like the one on the left" or "more like the one on the right," and so on. Subjects also made judgements with the length of B's path equal to 25, 35 and 40 mm.

This procedure was repeated with the other training stimulus, after which the whole procedure was repeated with the lengths of path reduced by similar steps from 40 to 5 mm.

In these training procedures, the steps employed were a little elastic, as the idea was to acquaint the subjects with the *nature* of the task rather than to give them practice at the more difficult task of the experiment proper.

Subjects then proceeded to the pilot experiment, which was similar in principle to the experiment to be reported in the following section, although different speed values were

used. The pilot experiment will not, therefore, be described, as its main function was to enable the experimenter to design the final experiment.

(1) *Main experiment*

Altogether 40 subjects took part in the present experiment. Of these, nine had taken part in the pilot experiment previously referred to. These, therefore, had received training in the task. Because we were, as a consequence, using fairly well-practised subjects, we thought it legitimate to employ also the well-trained laboratory personnel as subjects (investigation of their results showed later that these did not differ in significant respects from the results of other subjects). Besides the laboratory personnel and the subjects from the previous experiment, 26 new subjects took part.

The well-trained laboratory personnel proceeded directly to the experiment proper. Subjects from the previous experiment were again shown the prototype stimuli and asked to describe them, to ensure that no change of attitude had taken place since the last testing (this was found to be the case with three subjects, who were then eliminated from the experiment). They were then given brief practice to remind them of the task.

With the new subjects, the training procedure described in the previous section was carried out in full.

Subjects then proceeded to the experiment proper, and were instructed as follows (instructions were given in French to French-speaking subjects, and in English to the Americans who took part in the experiment.)

"I am going to show you a number of phenomena. In each case you will see the path of the red object very short at first, but I shall gradually increase it. Each time I do, I want you to tell me which of these two phenomena here [indicating the prototype stimuli] it resembles more, the one on the left or the one on the right. If it is like neither one nor the other, or if you are uncertain, say 'In between.' Do you understand?"

If the subject understood, he was told that he could look at the prototype stimuli as often as he wished, but was asked to make his decision quickly in each case, so that he would be reporting his first impression rather than his considered opinion.

In the experiment itself, the path of object B was increased in steps of 5 to 35 mm. each time. After the subject had been presented with all the stimuli, the experiment was repeated with the paths decreasing from 50 to 5 mm. in steps of 5 mm. In each case, the apparatus was covered from the subject's view when not in use and while the path lengths were being changed.

Each subject thus gave two responses for each trajectory length from 5 to 35 mm. inclusive. Since there were 40 subjects, there were altogether 80 responses at each trajectory length from 5 to 35 mm.

In order to balance out intra-serial practice effects, half the subjects were presented with the phenomena in one order, and half in the reverse order. The orders employed were as follows:—

Order 1	54:22·5	6:15	45:30	9:22·5	13·5:30	30:15
Order 2	30:15	13·5:30	9:22·5	45:30	6·0:15	54:22·5

It was appreciated that this did not control all sources of practice effects, but a design in which all these factors were controlled would have required an impossibly long time, not only on the part of the experimenter, but also of each subject (*vide infra.*).

Results

Table 1 shows the raw scores for the experiment. This Table should be read as two columns, the column on the left giving the results for a descending ratio, the column on the right for an ascending ratio. Under each ratio is given the number of responses of Launching (i.e. "more like the prototype stimulus 30·0:3·0) "L"; the number of responses of Triggering (i.e. "more like the prototype stimuls 4·5:30·0") "T"; and the number of "In between" responses, "I."

Examination of Table I reveals the following points.

First, in the case of a descending ratio of speeds, increasing the length of B's path appears to have no systematic effect on the tendency to report "Triggering"; there is a slight *decrease* in the tendency to report "Launching"; the number of "In between" responses shows a slight increase.

Second, in the case of an ascending ratio, increasing the length of the path is associated with a systematic increase in the number of reports of "Triggering" and a systematic decrease in the number of reports of "Launching." The number of "In between" responses fluctuates.

TABLE I

SHOWING THE CHANGE IN DISTRIBUTION OF RESPONSES OF LAUNCHING (L), TRIGGERING (T), AND "IN BETWEEN" (I), WITH CHANGE IN LENGTH OF B'S PATH

Descending ratios					*Ascending ratios*				
Stimulus 54:22·5 (*Ratio* 2·4:1)					*Stimulus* 6:15 (*Ratio* 1:2·5)				
Length of path	*Responses*				*Length of path*	*Responses*			
	L	*T*	*I*	*Total*		*L*	*T*	*I*	*Total*
5 mm.	51	23	6	80	5 mm.	37	35	8	80
10 mm.	54	19	7	80	10 mm.	31	34	15	80
15 mm.	51	19	10	80	15 mm.	23	42	15	80
20 mm.	50	16	14	80	20 mm.	21	47	12	80
25 mm.	45	17	18	80	25 mm	22	51	7	80
30 mm.	44	17	19	80	30 mm.	15	55	10	80
35 mm	43	22	15	80	35 mm.	13	56	11	80
Total	338	133	89	560	Total	162	320	78	560
Stimulus 45:30 (*Ratio* 1·5:1)					*Stimulus* 9:22·5 (*Ratio* 1:2·5)				
Length of path	*Responses*				*Length of path*	*Responses*			
	L	*T*	*I*	*Total*		*L*	*T*	*I*	*Total*
5 mm.	49	19	12	80	5 mm.	52	21	7	80
10 mm.	50	17	13	80	10 mm.	47	21	12	80
15 mm.	47	22	11	80	15 mm.	38	30	12	80
20 mm.	45	23	12	80	20 mm.	29	35	16	80
25 mm.	46	22	12	80	25 mm.	23	46	11	80
30 mm.	40	23	17	80	30 mm.	20	51	9	80
35 mm.	45	24	11	80	35 mm.	18	49	13	80
Total	322	140	88	560	Total	227	253	80	560
Stimulus 30:15 (*Ratio* 2:1)					*Stimulus* 13·5:30 (*Ratio* 1:2·25)				
Length of path	*Responses*				*Length of path*	*Responses*			
	L	*T*	*I*	*Total*		*L*	*T*	*I*	*Total*
5 mm.	60	12	8	80	5 mm.	48	18	14	80
10 mm.	63	10	7	80	10 mm.	35	29	16	80
15 mm.	64	9	7	80	15 mm.	30	32	18	80
20 mm.	61	10	9	80	20 mm.	16	45	19	80
25 mm.	59	15	6	80	25 mm.	12	52	16	80
30 mm.	56	11	13	80	30 mm.	7	61	12	80
35 mm.	51	14	15	80	35 mm.	11	56	13	80
Total	414	81	65	560	Total	159	293	108	560

The results for the two sets of ratios, combined and converted to percentages, are given in Table II. Since there were three ratios ascending and three descending, and since as pointed out there were 80 responses at each trajectory length for each

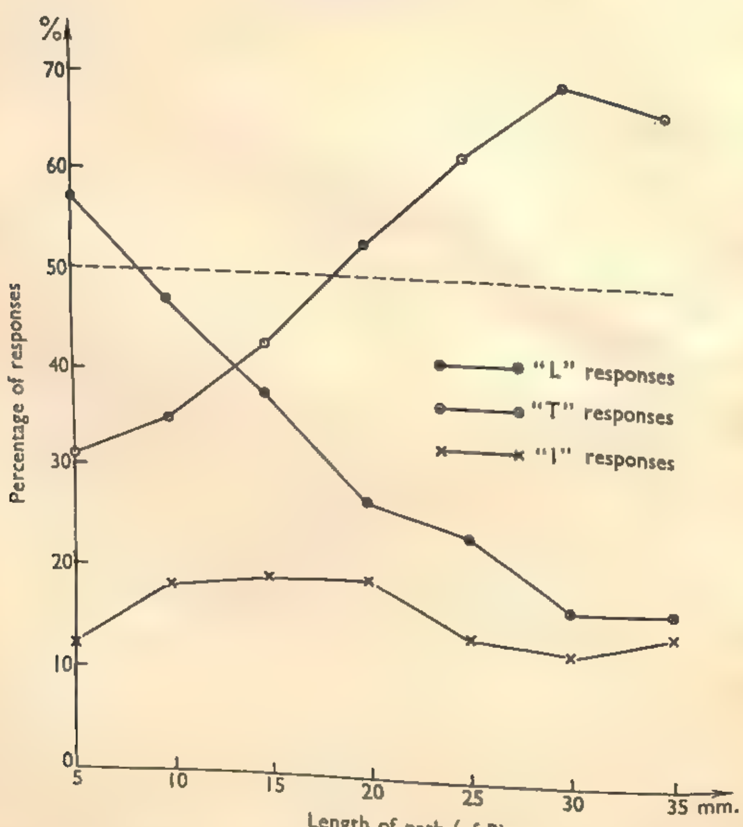
stimulus, these percentages are worked out on a base of 240. The results for ascending ratios are shown graphically in Figure 1.

TABLE II

DISTRIBUTION OF RESPONSES FOR EACH SET OF RATIOS, COMBINED AND CONVERTED TO PERCENTAGES

Descending ratios				*Ascending ratios*			
Length of path	*Percentages*			*Length of path*	*Percentages*		
	L	*T*	*I*		*L*	*T*	*I*
5 mm	66·5	22·5	11·0	5 mm.	57·0	31·0	12·0
10 mm.	69·5	19·0	11·5	10 mm.	47·0	35·0	18·0
15 mm.	67·5	21·0	11·5	15 mm.	38·0	43·0	19·0
20 mm.	65·0	20·0	14·5	20 mm.	27·5	53·0	19·5
25 mm.	62·5	22·5	15·0	25 mm.	24·0	62·0	14·0
30 mm.	58·5	21·0	20·5	30 mm.	17·5	69·5	13·0
35 mm.	58·0	25·0	17·0	35 mm.	17·5	67·0	15·5

FIGURE 1



Composite Results for Ascending Ratios, Expressed in Percentages (See Table II).

Discussion

The full implications of the present experiment cannot be appreciated without complete familiarity with Michotte's work, a discussion of which would be out of place in a paper of this nature. However, a number of observations can be made.

First, the results appear to confirm the experimental hypothesis. In the case of a descending ratio, increasing the length of B's path did not increase the tendency to report "Triggering." There was, as expected, a decrease in the number of reports of "Launching." The increase in the number of "In between" responses may be taken as evidence of the confusion brought about by the conflict of forces of organisation working in opposite directions.

In the case of an ascending ratio, the forces of organisation are working in the same direction, with the results that increase in the length of B's path is associated with an increase in the number of responses of "Triggering." In this case, the decrease in the number of "Launching" responses at greater lengths of B's path is much more marked. The combined results (Table II) show that the percentage of "In between" responses reaches a peak roughly in the same neighbourhood as the point where the curves for "Triggering" and "Launching" intersect. (See also Fig. 1.)

The relationships of the curves resemble those in the familiar judgement situation, where the response categories are, for example, "heavier than," "lighter than" and "equal to." It was in an attempt to establish this type of situation that subjects were presented with prototype stimuli and asked to compare the experimental stimuli with them.

However, there are certain dissimilarities between the conventional situation and our own that must be stressed. In the case of lifted weights, the responses are understood in the same way by all subjects, even if their actual judgements of which weights are heavier than, lighter than, or equal to the standard, differ. In the present case, among subjects who agreed in saying that a given experimental variable was "like" a prototype stimulus, the particular similarities found may have differed widely from one subject to another. To discover and define these differences one would have to question subjects extensively. We do not deny the value of the interview technique in experimental psychology, but in the present case this was irrelevant to our purpose, which was to demonstrate a relationship between changing stimulus conditions and a change in distribution of a limited number of responses.

It will be noted that, even with ascending ratios, the "Launching" response did not disappear entirely, even at the extreme length of path. It is not clear, however, whether the final value of about 17 per cent. represents an asymptotic value. In Michotte's opinion this was so, but it can be shown that the first six points fit a straight line. The seventh point, in the cases of "Triggering" and "In between" responses also, does not appear to fit the same curve as the earlier points of the graph.

A possibility here is that, after 30 mm., some other factor enters into the situation. This factor will, very probably, be the radius of action which as was mentioned in the introduction will limit the length of path in an experiment of this sort. 30 mm. appears to be the upper limit in this case. It is possible that the "Launching" responses will disappear entirely, if the length of path is increased beyond this point, but more likely that a new set of forces of organisation will come into operation. However, more experimentation is necessary before a decision can be made on this point.

It was mentioned in the Section on Procedure that about 50 per cent. of the subjects had to be rejected on the basis of the preliminary screening test, either because their replies were completely ambiguous or because they described the

prototype Launching disc in terms appropriate to Triggering, and the prototype Triggering disc in terms appropriate to Launching. The reversal of the expected responses was commoner than the ambiguity, and appears to conflict with Michotte's reports of the near universality of agreement of response. However, it must be pointed out that only two ratios were used in the prototype stimuli, and that these may not have been the most apropriate for establishing the impressions. The case of A:B = 30·0:3·0 in particular is likely to be disturbing, as the most suitable ratio for the impression of Launching is about 3 or 4:1. Consequently, the particular prototype stimuli used may have tended to cause ambiguity of impression, and the response would have depended on the subject's perceptual attitude. This is supported by the case of the three subjects in the Main Experiment (mentioned in the section on Procedure), who responded differently on different occasions. However, this question could be the subject of an inquiry in its own right, and we mention it only to warn against the danger of drawing conclusions from this particular observation.

The present experiment differed in an important way from Michotte's experiments. As has been pointed out, Michotte asked subjects to describe their "impressions," and put no further restrictions on the response categories. In the major part of the present investigation, subjects were asked to compare the experimental stimuli with the prototype stimuli. Critics may object to this modification on the grounds that too much restriction is in this way imposed upon the data. This objection would be part of a much wider controversy in experimental psychology, into which it is impossible to enter here. Suffice it to say that one has the choice of extremely broad response categories, which permit one to uncover as much as possible of the subject's attitude; or narrow response categories that tell the experimenter little in themselves but which may be interpreted with the minimum of ambiguity. In order to achieve the "richness" of data that the former method permits, one has to increase the number of response categories. It is the present experimenter's opinion that the former method is more suitable for preliminary research, the latter for further experimentation which serves to make the research findings more precise.

In spite of the attempt to make the responses more precise, by limiting the scope and nature of the response categories in the present experiment, there was still room for ambiguity. Later experimenters may reduce this ambiguity yet further. Further experimentation should also be concerned to control a number of variables not controlled in the present case. *First*, the effects of order of presentation may have an important effect on the results. The ideal would have been to present the different path lengths in random order but, to do this in such a way that every path length was presented in the same relation to every other path length would have taken more time than was available to the experimenter, and more time than any subject would have been prepared to give. A re-design of the experiment, using the technique of analysis of variance, may help to solve the problem of prohibitive time. *Second*, although the results of well-trained subjects did not differ significantly from those of subjects less well-trained in the present case, this does not mean that practice has no effect upon the judgements. Training of new subjects was carried out in the present case in order to teach subjects the experimental response vocabulary; however, subjects who had been accustomed to this type of experimentation for years would certainly have approached the task with a different attitude from those experiencing it for the first time. In particular, they would already have their concepts of "Launching" and "Triggering" firmly established, and would respond without direct reference to the prototype stimuli. The present experiment was not sufficiently sensitive to detect any differences in the responses, but it is quite possible that

experiments designed to take account of the effect of length of experience would be able to measure this effect, at least in a rudimentary form.

There is another, more specific criticism that may be raised against the present design. In the experiment, subjects were permitted a response of "In between" when they could not be sure which of the prototype stimuli the experimental stimulus resembled more. The controversy about the permissibility of an "uncertain" category, and the advisability of employing a forced-choice choice technique is of long standing in the history of experimental psychology, and we cannot, in this paper, justify our choice of technique at any length. Briefly, we permitted the third category because, in Michotte's opinion, if subjects did not regard the experimental stimulus as unambiguously resembling either of the prototype stimuli, the subjective impression may have been one of many not yet precisely defined. The third category, therefore, was a "blanket" category, to be broken down by future inquiry.

One final point on which future experiment seems to be required is the question of individual differences. Besides the differences, already referred to, with regard to the possible different effect of different ratios on the impressions, the whole question of exactly what the stimulus resembles, and how the resemblance may be changed by a change of perceptual attitude, requires extensive further research.

References

Michotte A. (1954). *La Perception de la Causalité.* Ed. II. Louvain.
Vernon, M.D. (1952). *A Further Study of Visual Perception.* London.
Yela, M. (1954). La nature du "Rayon d'Action" dans l'impression de causalité méchanique. *J. Psychol. norm. path.* **47-51**, 330–48.

SHORTER ARTICLES AND NOTES

GROUPING STRATEGIES WITH SIMULTANEOUS STIMULI

BY

J. A. GRAY and A. A. I. WEDDERBURN

From the Institute of Experimental Psychology, University of Oxford

Broadbent (1956) reports that two lists of digits, each presented to one ear separately so that the items in the two series coincide in time, are grouped together according to the ear-of-arrival, and that these two lists are accordingly recalled separately, one after the other. To ascertain whether such a tendency reflected some built-in mechanism or whether it was due to an optional tactic, adopted through success in making sense of message sequences in other situations, an experiment was designed in which a meaningful message would emerge for the subject if the ear-of-arrival cue was *ignored*. In this experiment, words broken up into syllables, and phrases broken up into their monosyllabic constituent words were presented to the subject, with the constituents alternating between the two ears. At the same time lists of digits were presented to whichever ear was unoccupied. The results show that recall by meaning rather than by ear-of-arrival, when these are in conflict, can occur and is no less efficient.

Introduction

Broadbent (1956 and 1958) has reported that when two lists of three digits are presented simultaneously to separate ears at a speed of two digits a second on each ear, "in the vast majority of cases . . . all the information from one channel appeared in response before any of the information from the other channel" (p. 212, 1958). This finding is peculiarly interesting because grouping by temporal order, which is the normal phenomenon at slow speeds, seemed to be unobtainable. However his material, single digits, offered no further cues for grouping other than the ear on which each item arrived. It is not clear, therefore, whether the preference shown by subjects for grouping in this way reflects some built-in mechanism, or is simply an optimal strategy for the situation—i.e. a bias on the basis of past success in classifying the input of one ear as a continuous message, when other coincident cues (e.g. message sense, voice, pitch, etc.) also fit into this classification.

If grouping-by-ear is an acquired bias, the presence of other cues *in conflict with* the ear of arrival should be able to disrupt it, and some other strategy of recall will be used.

In the following experiments, such cues were provided by (1) 3-syllable words broken up into syllables and presented to alternate ears, and (2) 3-word phrases used in the same way.

Experiment I

Equipment

A Brenell Mark IV 2-channel tape-recorder, with each channel feeding into a separate headphone, was used throughout. Subjects recorded their own responses on cards at the end of each list.

Subjects

Ten male and 5 female university undergraduates, naïve in dichotic listening situations, were given preliminary practice with one series of digits only, under the same overall conditions as in the subsequent experiment.

Material

Ten lists of digits and syllables, arranged as in the following examples, orders (1) and (2) alternating throughout:—

(1) Left	Right	(2) Left	Right
EX-	6	1	CYC-
2	TIR-	LO-	7
PATE	9	3	STYLE

The other words used were: fol-i-o, le-pre-chaun, ob-jec-tive, mas-ter-ful, se-man-tic, cad-mi-um, ri-vu-let, hy-dro-gen. The digits were taken from a table of random digits, though no list contained any digit twice.

Procedure

The interval between the onset of each simultaneous pair was half-sec., and the lists were 13 sec. apart, which gave ample time for subjects to record their results.

Subjects were divided into two groups. Group 1 (7 subjects) was given neutral instructions, to expect syllables and digits, and to write down exactly what was heard at the end of each list. They were not told that the syllables formed words. Subjects in Group 2 (8 subjects) were told in addition that the syllables formed three-syllable English words, and were instructed to listen for these words as a unit, as well as the three digits. Both Groups were asked to write down what they recalled immediately at the end of each list, and not to "reshuffle" the items on paper, nor, as far as possible, in their heads. The second Group was also asked not to attempt to infer or reconstruct a word if they had failed to hear it as a word. Subjects were watched throughout to ensure that they did not re-order the items as they wrote them.

Results

The results are given in Table I. The low mean scores indicate that all subjects found the task difficult, except Group 1 with the digits, and overall comparison of the grouping used was impossible owing to the number of ambiguous responses containing gaps and reduplicated items.

TABLE I

(*a*) Per Cent. Items Recalled

	Digits	*Syllables*	*Mean*
Group 1	78	49	64
Group 2	55	36	46

(*b*) Grouping of All-correct Lists

	Ear	*Meaning*	*Indeterminate*	*Total*
Group 1 ..	4(1.S)	—	1	5/70
Group 2 ..	1	8 (4.S)	—	9/80

But despite their higher score on individual syllables, Group 1 never made a whole word, and only had 5 lists with all 6 items correct, 4 being grouped by ear, and only the fifth one being free of phonetic approximations. Of the 9 all-correct lists of Group 2, 7 were perfectly recalled, and one had the syllables collected with a slight distortion, (CAT-ME-UM for CAD-MI-UM) without the word being recognized. Here plainly transition probabilities between syllables were not being utilized, but the differentiation between syllables and digits was alone sufficiently strong as a cue to break down the tendency to group-by-ear.

Two of the words ("leprechaun" and "rivulet") were never reported at all, perhaps owing to unfamiliarity.

We can conclude from this experiment that Group 2, owing to the greater difficulty of the task set to them, showed lower overall efficiency (fewer *items* recalled), but were more completely successful (more all-correct *lists*) when they achieved the technique of grouping-by-meaning.

Experiment 2

The second experiment was done by all subjects, in the same two Groups, immediately after Experiment 1.

Material

Ten lists of words and digits, arranged as in the following examples, orders (1) and (2) alternating throughout:—

	(1) *Left*	*Right*	(2) *Left*	*Right*
	Mice	3	2	who
	5	eat	goes	3
	cheese	4	9	there

The other phrases used were: What the Hell, did you see, my old flame, it's in there, there she goes, tea is laid, not my sort, dear Aunt Jane.

Procedure

The timing and instructions were, *mutatis mutandis*, the same as in Experiment 1. Group 1 were told that they would hear digits and words, Group 2 that they would hear digits and words forming 3-word phrases or sentences.

Results

Higher means (Table III*a*) permit a more complete analysis of the data. First, each list was categorized for scoring as "Ear," "Meaning," or "Indeterminate" according to strict conventions. Cases where mixed strategies occurred were allotted to "Indeterminate," as were lists reported with ambiguous gaps, guesses, or reduplications. (One subject specialized in mixed strategies, e.g. "Tea 3 1 is laid 7," achieving 3 all-correct lists and the second highest total overall score in Group 1. It is noteworthy that this is exactly the strategy of temporal sequence that Broadbent's subjects found impossible.)

TABLE II

Efficiency of Performance in Relation to Method of Grouping

Method of grouping	*Number of subjects*	*Per cent total lists so grouped*	*Per cent. correct items*		
			Digits	*Words*	*Mean*
Group 1 (7 Subjects)					
Ear	4	21	80	56	68
Meaning	6	37	74	65	70
Indeterminate	—	42			
Group 2 (8 Subjects)					
Ear	5	18	71	60	66
Meaning	8	56	81	93	87
Indeterminate	—	26			

Column 2 shows the preferred method of subjects; some subjects used both methods. Column 3 shows the proportion of the total *lists* in each category. The last three columns give the percentage of individual *items* correctly recalled in each method; this provides a measure of the relative efficiency of the two methods.

Table II gives the percentage lists clearly grouped-by-ear and grouped-by-meaning, and the number of subjects using each of these methods. The *percentage success* of

grouping-by-ear versus grouping-by-meaning was computed by calculating the percentage of individual items correctly reported out of the total possible within these categories; this is given for digits and words separately. It can be seen from this table that both Groups preferred grouping-by-meaning, though most subjects used both methods, and that both Groups (especially Group 2) were more successful when using this preferred method. This suggests that subjects were acting intelligently in the situation.

TABLE III

(*a*) Per Cent. Items Recalled (including "Intermediate")

	Digits	*Words*	*Mean*
Group 1	72	53	63
Group 2	72	74	73

(*b*) Grouping of All-correct Lists

	Ear	*Meaning*	*Indeterminate*	*Total*
Group 1 ..	5 (1.S)	1	4 (1.S)	10/70
Group 2 ..	1	21 (8.S)	—	22/80

Finally in Table III, the two Groups are compared for overall efficiency (including here the "Indeterminate" category), and for the number of all-correct lists, and their method of grouping. The overall efficiency for Group 1 is not significantly different from their score in Experiment 1; for Group 2 it is much higher on both digits and words. Every subject in Group 2 achieved at least one all-correct list in the second experiment.

Discussion of Results

It is clear from these results that the tendency to group simultaneously presented dichotic signals according to the ear to which they are presented can be overcome, given sufficiently strong cues favouring some other mode of grouping. This is apparent in Experiment 1, and clearly established in Experiment 2, even in the Group without instructions to follow meaning.

This system of grouping is if anything more efficient in this situation (see Tables I(*b*), II, and III(*b*)). If then we assume that the situation should require "switching attention," two possible alternatives have been suggested by Broadbent. (1) "It is possible to take a more passive attitude, and wait till after the stimuli have all arrived before beginning to deal with them successively," but "such an attitude is inefficient" (Broadbent 1957). (2) When some of his subjects did achieve immediate alternation of channels to a limited extent, he suggests "that it was done by a preliminary response to the information in some other order followed by a transposition of this order into the required one before the digits were reproduced publicly. Such a transposition is known from intelligence testing to be possible though more difficult than normal memory span" (Broadbent, 1958, p. 212). If we equate "more difficult" with "less efficient," this alternative is also ruled out here. So either our subjects were "switching," or the suggested alternatives are wrongly judged to be less efficient; or the "switching of attention" is not necessarily involved at all.

A further possibility is that subjects were utilizing transition probabilities between syllables and between words to guess at the third member of the word or phrase. We can distinguish between two ways in which this could happen. If subjects were guessing at the syllable or word they heard *last*, they must already have heard the first two syllables/words, as meaningful units, although this involved signals arriving at opposite ears. In this case, then, grouping-by-ear has already broken down before guessing can take place.

If, secondly, they were guessing at the middle syllable/word, this need not involve attention to both ears at all, as the first and last meaningful units always arrived on the same ear. However this would not account for those cases where subjects successfully recalled the *three digits* as well in the required order, for there was no way of inferring from one to another of these.

Moreover we should expect guessing to show itself in *semantic* substitutes for the correct word/syllable. Though this did occur five times in all, there were more substitutions of a word/syllable *phonetically* similar to the correct signal (nine cases), e.g. "not my salt" for "sort"—showing that the correct signal had been heard, even if imperfectly.

A final point requiring explanation is the greater difficulty for Group 2 of the task set in Experiment 1 as compared with Experiment 2. The difference is probably due to the fact that words are heard from the first as meaningful units, whereas syllables only become meaningful when the whole word is known, and are therefore less useful as immediate cues-for-grouping.

Conclusions

In terms of our original question, these results suggest that the ear-of-arrival is only one possible cue-for-grouping, although it is often dominant, especially in otherwise undifferentiated messages, as was the case for Broadbent's subjects. But by biassing the material so that meaning-cues must be followed for maximum pay-off, this skill can be displaced; the odd idiosyncratic variations, not reflected in the quantitative results, support our view that subjects are searching for an optimal strategy, and if helped by instructions, readily adopt it. Subjects are simply using whatever cues are available to interpret sensory events.

We acknowledge our thanks to Dr. J. A. Deutsch for suggesting the experiment and for supervision and many helpful criticisms, to Neville Moray for technical advice, encouragement and discussion, and to Ann Taylor for the patient and generous loan of valuable equipment and space. We also thank Professor Oldfield for providing space and facilities for the experiment.

References

Broadbent, D. E. (1956). Successive responses to simultaneous stimuli. *Quart. J. exp. Psychol.*, **8**, 145–52.

Broadbent, D. E. (1957). Immediate memory and simultaneous stimuli. *Quart. J. exp. Psychol.*, **9**, 1–11.

Broadbent, D. E. (1958). *Perception and Communication.* London.

COMBINATION OF DRIVE AND INCENTIVE*

BY

IRENE MACKINTOSH HULICKA
From the Veterans Administration Hospital, Buffalo, N.Y.

This experiment was designed to investigate the combination of drive and incentive as determinants of performance. Nine groups of rats were trained to press a bar under three levels of food deprivation (12, 24 and 36 hr.) and three incentive conditions (1, 2 and 3 pellets). Response strength was estimated by counting the number of responses with a latency of 1 sec. or less during five 20 trial sessions. The results indicated that response strength increased with hours of deprivation and with amount of food reward. Significant interactions between sessions and drive, and sessions and incentive provided additional support for the multiplicative combination of H (habit) and D (drive) and H and K (incentive). The lack of significant interaction between D and K was interpreted as supporting the hypothesis that D and K combine additively rather than multiplicatively.

This experiment was designed to investigate the manner in which the theoretical motivational variables drive (D) and incentive (K) combine with habit (H) to determine response strength (E). Evidence for the multiplicative interaction of H and D is provided by several investigations (Passey, 1948; Perin, 1942; Ramond, 1954; Williams, 1938) in which the number of reinforced trials and hours of deprivation or intensity of the unconditioned stimulus were varied. Response measures, including frequency, latency and number of responses to extinction, plotted against the number of trials show with one exception (Campbell and Kraeling, 1953) diverging curves for the different drive groups. Consequently the bulk of the evidence indicates that $E = f(H \times D)$. When D is held constant but amount or delay of reward is varied and a response measure is plotted against number of trials diverging curves are also obtained (Crespi, 1942; Dufort and Kimble, 1956; Fletcher, 1940; Grindley, 1929; Heyman, 1957; Perin, 1943) indicating that $E = f(H \times K)$. These conclusions require the assumption made by Hull (1951, 1952) and others that H is a function of the number of reinforced trials but is not affected by variations in D or K. Although there is some evidence against this assumption (Eisman, Asimow and Maltzman, 1956), more evidence is in support of it (Brown, 1956, Deese and Carpenter, 1951; Heyman, 1957; Hillman, Hunter and Kimble, 1953; Kendler, 1945; Strassburger, 1950; Teel, 1952). The limited generalization that $E = fH(D, K)$ seems permissible. Lacking experimental evidence concerning the combination of D and K, Hull (1943, 1951, 1952) speculated that they combined in a multiplicative manner. Reynolds, Marx and Henderson (1952) using resistance to extinction as a measure of response strength obtained evidence in favour of Hull's hypothesis. Recently, Spence, (1954, 1956) suggested that D and K might combine additively. An experiment by Loess (1952) may be interpreted as favouring this suggestion. Sufficient data are not available to evaluate the alternative hypotheses.

This experiment was designed to evaluate the additive and multiplicative hypotheses for the combination of D and K. Each of drive (hours of deprivation) and incentive (pellets of food) were varied three ways in a simple learning situation. If the effects of D and K on performance are unrelated, the additive hypothesis would be supported. If there is an interaction between the effects of D and K the multiplicative hypothesis would be supported.

* This research was supported by a grant from the National Science Foundation.

Method

Subjects and apparatus. The subjects, 90 rats purchased from the Holtsman Company, were approximately 100 days old at the beginning of the experiment. For 8 days prior to the experiment, all animals were kept on a reduced diet of 8 gm. of Purina chow per day. The apparatus, a modified Skinner box, consisted of two compartments, an entrance box and a response box. The response box was equipped with one bar, which automatically withdrew after a bar press response. An electric timer was activated when the experimenter opened the door connecting the two boxes and stopped when the animal pressed the bar.

Procedure. Prior to the introduction of the independent variables, D and K, all animals were given 50 reinforced bar press trials. All were deprived of food for 24 hr. before each of these experimental sessions. The reward for each bar press response was two

TABLE I

Experimental Design

Hours of deprivation	*Number of Pellets*		
	1 (K1)	2 (K2)	3 (K3)
12 (D12)	Group 12.1	Group 12.2	Group 12.3
24 (D24)	Group 24.1	Group 24.2	Gtoup 24.3
36 (D36)	Group 36.1	Group 36.2	Group 36.3

TABLE II

Means and S.D.'s for Frequency of Responses with Latency of 1 sec. or Less during 100 Trials

Group		*K1*	*K2*	*K3*	*Total for D Groups*
D12	Mean S.D.	30·4 5·5	38·4 5·2	42·8 7·9	37·2 8·2
D24	Mean S.D.	37·7 5·5	41·6 6·7	46·5 5·7	41·6 8·8
D36	Mean S.D.	41·6 7·7	46·4 6·8	48·3 4·5	45·4 7·3
Total for K groups	Mean S.D.	36·6 7·8	42·1 7·3	45·9 6·4	41·5 8·3

45 mgm. pellets. On each trial the animal was placed in the entrance box, entered the response box, and remained there for 40 sec. after the bar press response, sufficient time for the consumption of the pellets. The interval between trials was approximately 15 min. Four trials were given on the first day and 12 trials per day were given for the next 4 days.

After the five days of preliminary training the 90 rats were assigned to 9 groups of 10 each. The groups were equated on the basis of response latency during the final 24 trials of preliminary training. Training was continued with variations in hours of deprivation (D) and pellets (K) as indicated in Table I. During this phase of the experiment appropriate reductions were made in the amount of food given in the home cage to account for the number of pellets given in the experimental situation. Excluding these reductions, all groups were fed 8 gm. immediately after each experimental session. Twelve, 24 or 36 hr. preceding the next experimental session, another 8 gm. were supplied. After the introduction of D and K variations, each animal was given 20 spaced trials every other day for 5 days. The latency of each bar press response was recorded.

Results

The results were analyzed in terms of the frequency of trials for which the response latency was 1 sec. or less. Table II presents the means and S.D. for this measure of response strength for the 100 trials after the introduction of D and K variations.

TABLE III

Analysis of Variance of Frequency of Responses with Latency of 1 sec. or Less

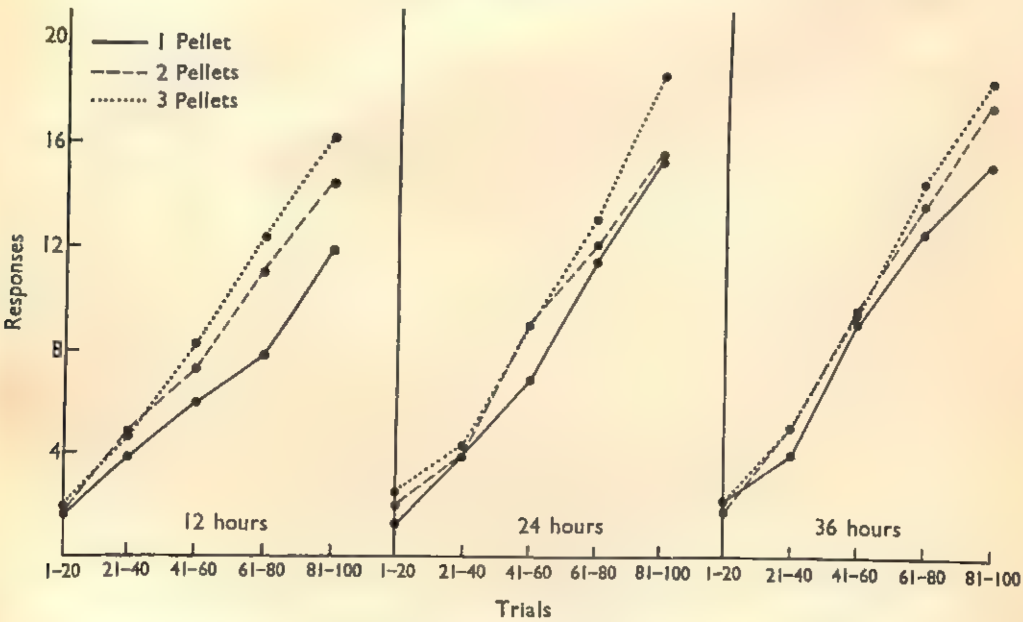
Source	*d.f.*	*Mean Square*	*F*	*P*
Between drives	2	512	11·58	<0·01
Between incentives	2	657	14·86	<0·01
Incentive × drive	4	26	0·59	>0·05
Within	81	44·2	—	—

Table IV

Analysis of Variance of Frequency per 20 Trials of Responses with Latency of 1 sec. or Less

Source	*d.f.*	*Mean Square*	*F*	*P*
Between drives	2	1024	30·1	<0·01
Between incentives	2	1314	38·5	<0·01
Between sessions (20 trials)..	4	27982	824·0	<0·01
Drive × incentive	4	52	1·5	>0·05
Drive × sessions	8	160	4·5	<0·01
Incentive × sessions.. ..	8	144	4·2	<0·01
Drive × incentive × session	16	34	—	—

FIGURE 1



The mean frequency per 20 trials of responses with a latency of 1 sec. or less for each of 9 experimental groups.

Table III shows the results of analysis of variance when applied to these data. Figure 1 summarizes for each of the 9 groups the mean frequency per 20 trials of responses with latency of 1 sec. or less. In Figure 1 the three separate graphs represent the three drive conditions. A similar graph could be prepared to represent the three incentive conditions by rearranging the curves. The results of analysis of variance applied to the data in Figure 1 are given in Table IV. A separate analysis was done to determine whether the differences between drive groups and incentive groups during the first 20 trials were significant. The obtained F's did not approach statistical significance.

Discussion

The results indicated, as would be expected on the basis of other investigations, that response strength as measured by the frequency of responses with a latency of 1 sec. or less, was a function of trials, hours of deprivation and amount of reward. Since the bar press response was not well established prior to the introduction of the D and K variations, the differences between experimental sessions reflects the growth of H. The numerous investigations which indicate that H is not a function of D or K make it permissible to assume that H increased equally for all groups. Therefore, the significant interactions between incentives and sessions and between drives and sessions, and the diverging curves shown in Figure 1 supply additional evidence in favour of the multiplicative combination of H and K and of H and D.

The purpose of this experiment was to provide information concerning the combination of D and K. If D and K combine multiplicatively, the differences between K groups should vary with D and vice versa. If D and K combine additively, differences between K groups should not vary with D and vice versa.

For example,

If $E = fH(D \times K)$	If $E = fH(D + K)$
$E\,12.2 - E\,12.1 = HD\,12\,(K2 - K1)$	$E\,12.2 - E\,12.1 = H(K2 - K1)$
$E\,24.2 - E\,24.1 = HD\,24\,(K2 - K1)$	$E\,24.2 - E\,24.1 = H\,(K2 - K1)$

Thus if $E = fH(D \times K)$ the interaction between D and K should be significant. The fact that the obtained interaction is not significant may be interpreted as evidence against the multiplicative and for the additive hypothesis for the combination of D and K.

These results are somewhat at variance with those reported by Reynolds, *et al.* (1952). They found that during training their high drive—high reward group (our Group 36.3) exhibited longer response latencies than the low drive—high reward group (our Group 12.3). Our results are not in agreement with this portion of their findings. They, however, reported that their low drive—low reward (our Group 12.1) exhibited longer response latencies than the high drive—low reward group (our Group 36.1). Here, our results correspond to theirs. One reason for the different results may be that their high drive group was deprived of food for 48 hr. while ours was deprived for 36 hr. It is possible, as Spence (1956) suggests, that at the extremely high drive level, competing responses are more apt to occur.

During preliminary training all animals were given 2 pellets per bar press. For the remainder of the experiment, the animals were rewarded with 1, 2, or 3 pellets respectively. No significant difference was found between incentive groups during the first 20 trials after the change in reward. Thus, this experiment provides no evidence for abrupt changes in performance following reward switching. This finding is in agreement with that of Pereboom (1957) and provides additional evidence for his suggestion that Hull's Theorem 30 requires revision.

Although the results of this experiment may be interpreted as indicating that D and K combine in an additive manner, the limited generalizability of this conclusion should be recognized. Before definitive general statements can be made about the combination of D and K variables, a variety and range of drives and incentives should be used, with different responses, more complex learning situations, and different levels of H.

References

Brown, J. (1956). The effect of drive on learning with secondary reinforcement. *J. comp. physiol. Psychol.*, **49**, 254–60.

Campbell, B. A., and Kraeling, D. (1953). Response strength as a function of drive level and amount of drive reduction. *J. exp. Psychol.*, **45**, 97–101.

Crespi, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, **55**, 467–517.

Deese, J., and Carpenter, J. A. (1951). Drive level and reinforcement. *J. exp. Psychol.*, **42**, 236–8.

Dufort, R. H., and Kimble, G. A. (1956). Changes in response strength with changes in the amount of reinforcement. *J. exp. Psychol.*, **51**, 185–91.

Eisman, F., Asimow, A. and Maltzman, I. (1956). Habit strength as a function of drive in a brightness discrimination problem. *J. exp. Psychol.*, **52**, 58–64.

Fletcher, F. M. (1940). Effects of quantitative variation of food incentive on the performance of physical work by chimpanzees. *Comp. Psychol. Monogr.*, **16**, No. 82.

Grindley, G. C. (1929). Experiments on the influence of amount of reward on learning in young chickens. *Brit. J. Psychol.*, **20**, 173–80.

Heyman, W. (1957). Certain relationships between stimulus intensity and stimulus generalization. *J. exp. Psychol.*, **53**, 239–48.

Hillman, B., Hunter, W., and Kimble, C. (1953). The effect of drive level on the maze performance of the white rat. *J. comp. physiol. Psychol.*, **46**, 87–9.

Hull, C. L. (1943). *Principles of Behavior*. New York.

Hull, C. L. (1951). *Essentials of Behavior*. New Haven.

Hull, C. L. (1952). *A Behavior System*. New Haven.

Kendler, H. H. (1945). Drive interaction: II. Experimental analysis of the role of drive in learning theory. *J. exp. Psychol.*, **35**, 188–98.

Loess, H. B. (1952). The effect of variations of motivational level and changes in motivational level on performance in learning. *Unpublished doctoral dissertation, State Univ. of Iowa.*

Passey, G. E. (1948). The influence of intensity of unconditional stimulus upon acquisition of a conditional response. *J. exp. Psychol.*, **38**, 420–8.

Pereboom, A. C. (1957). An analysis and revision of Hull's Theorem 30. *J. exp. Psychol.*, **53**, 234–7.

Perin, C. T. (1942). Behavior potentiality as a joint function of the amount of training and degree of hunger at the time of extinction. *J. exp. Psychol.*, **30**, 93–113.

Perin, C. T. (1943). A quantitative investigation of the delay of reinforcement gradient. *J. exp. Psychol.*, **32**, 37–51.

Ramond, C. K. (1954). Performance in selective learning as a function of hunger. *J. exp. Psychol.*, **48**, 265–70.

Reynolds, B., Marx, M., and Henderson, R. (1952). Resistance to extinction as a function of drive-reward interaction. *J. comp. physiol. Psychol.*, **45**, 36–42.

Spence, K. W. (1954). Current interpretations of learning data and some recent developments in stimulus response theory. In *Learning Theory, Personality Theory and Clinical Research.* New York.

Spence, K. W. *Behavior Theory and Conditioning.* New Haven.

Strassburger, R. C. (1950). Resistance to extinction of a conditioned operant as related to drive level at reinforcement. *J. exp. Psychol.*, **40**, 473–87.

Teel, K. S. (1952). Habit strength as a function of motivation during learning. *J. comp. physiol. Psychol.*, **45**, 188–91.

Williams, S. B. (1938). Resistance to extinction as a function of the number of reinforcements. *J. exp. Psychol.*, **23**, 506–22.

BOOK REVIEWS

Effect of Small Doses of Alcohol on a Skill Resembling Driving. By G. C. Drew, W. P. Colquhoun and Hazel A Long. Medical Research Council Memorandum No. 38. London. H.M.S.O. 1959. Pp. 108. 7s. 6d. net.

The purpose of this inquiry was to investigate the effects of small doses of alcohol upon sustained performance on the "Miles Motor Driving Trainer"—a complex tracking task presumed to have a good deal in common with driving. Forty subjects, average age 31 years and all with driving experience, took part in the experiments, which extended over five separate 2-hr. sessions. On each session the subjects received one of four alcohol doses or a placebo dose, the highest alcohol dose being roughly equivalent to 5 fluid ounces of whisky for a man of average weight. Subjects were not, however, informed as to which of the five types of dose they had received. The condition of the stomach was kept as constant as possible and blood, urine and breath samples were taken at half-hourly intervals. A series of personality tests was also administered.

The principal results were as follows: (1) Mean error increased significantly with increasing blood alcohol; with a blood alcohol concentration of some 80 mg. per cent. the deterioration in accuracy amounted to about 12 per cent. (2) The shape of the mean error curve was similar to that of the blood and urine alcohol curves. (3) The average speed of "driving" showed no significant change with increasing blood alcohol. (4) Individual differences were rather marked and personality assessments, especially those relating to introversion-extraversion, showed a definite relation to the behaviour changes.

Professor Drew and his colleagues comment sensibly on these findings and show exemplary caution in arguing from them to driving under ordinary conditions. They provide a most useful analysis of changes in the various components of performance (which may or may not lead to decrement in overall achievement) and of the incidence and nature of compensatory activities. They also lay proper stress on temperamental factors in relation to individual differences in reaction to alcohol.

It is obvious that the results of this study have close—perhaps decisive—relevance to a vexed social issue. At the same time, it is far from easy to translate the findings into real-life equivalents. In the first place, subjects were not told whether they had been given alcohol (and if so, how much) and for this reason might be expected to perform with somewhat excessive caution. This might affect particularly the results obtained under placebo conditions. In the second place, the experimental situation clearly provided little or no opportunity for the exercise of judgement or decision, which may well prove of far greater significance than formal tracking-ability in regard to the effects of alcohol upon driving skill. And in the third place, it might seem a pity that reliance in assessing individual differences was placed almost wholly upon rating scales and personality inventories, the limitations of which are widely acknowledged. Although we may welcome this intrusion of personality study into the analysis of skill, one may hope that it will develop along lines somewhat less constricted and more realistic than those advocated here.

This is a competent and level-headed report in the tradition of W. H. R. Rivers and H. M. Vernon. Sophisticated in design, the experiments measure up to the highest standards of contemporary psychology. The authors have taken full advantage of the advances in method that grew out of the war-time work of the Cambridge Psychological Laboratory and have applied to good effect the newer notions of skill and its deterioration which we owe above all to Sir Frederic Bartlett. The next step, perhaps, is to attempt to envisage these processes in concrete neurological terms.

O. L. Zangwill.

The Anatomy of Judgment. By M. L. Johnson Abercrombie. London. Hutchinson. 1960. Pp. 156. 25s. net.

This is a most interesting little book, written by a zoologist, which describes an experimental attempt to evoke freer and more fruitful ways of observing and thinking among young medical students. In a series of discussion classes, Mrs. Abercrombie started a number of likely hares (such as the meaning of the word "normal" in an anatomical text, or the facts and implications to be gleaned from a pair of radiographs of the child and adult human hand). The resulting wrangles bore some of the marks of a group

analytic situation and involved feelings and attitudes not always comfortable or immediately satisfying to the participants. A mass of fascinating data interesting to the unprejudiced student of observation, thinking and judgement was thrown up, which the author scrutinizes with imagination and insight within a framework of contemporary psychological concepts. She succeeds, without labouring theoretical interpretations, in making skilful and effective use of ideas deriving from Ames, Bartlett and Foulkes, though it is disappointing that she does not press her analysis a little nearer to the essence of judgement and decision. Anybody concerned with university teaching will recognize that the fresh ground she breaks will have, before long, to be much more deeply cultivated. It offers, as Mrs. Abercrombie has shown, a grand opportunity for the concrete study of the "higher mental processes" in genesis and action. Any teacher in whom this book arouses disquiet will gain from it. Psychologists who read it will find double cause for gratitude—and perhaps a spur to go and do likewise.

R. C. Oldfield.

Handbook of Physiology. Section 1*: Neurophysiology. Volume* 1. Editor-in-Chief: John Field. Section Editor: H. W. Magoun. Executive Editor: Victor E. Hall. American Physiological Society (Sole Agents: The Williams & Wilkins Co., Baltimore). Pp. xiii + 779. 1959. $22.00.

This monumental venture on the part of the American Physiological Society aims to provide a systematic and authoritative compendium of present-day physiological knowledge. The first section is devoted to neurophysiology and is to comprise three volumes, of which this is the first. It is understood that the remaining two volumes are in active preparation and are due to appear in the near future.

The present volume consists of thirty odd chapters loosely grouped in four parts. The first part is concerned with problems of neuron physiology, nervous conduction and neuromuscular transmission and is fittingly introduced by Prof. J. C. Eccles. The papers in the second part deal with brain potentials, intrinsic rhythms, evoked responses and some aspects of epilepsy; this section is introduced by Prof. A. Fessard. Lord Adrian leads off the third section, concerned with sensory mechanisms and the special senses; his short introductory essay is a model of relevance and clarity. The final section is wholly devoted to vision and is ably introduced by Prof. H. K. Hartline. Mention should also be made of the erudite opening chapter by Dr. Mary Brazier on the historical development of neurophysiology.

It will be denied by no one that Dr. Magoun has chosen his team of contributors extremely well. In general, the authors succeed admirably in the difficult task allotted them and only in rare cases does critical judgment appear deficient. If fault there be, it is perhaps shown in a certain tendency to overstress work too recent for it to be placed in proper perspective. This is correlated with the omission of much earlier work still accepted as standard. Although neurophysiology is advancing so fast that a correct perspective is often difficult, a more conservative attitude might well have resulted in a better balanced text. After all, a Handbook is—or should be—an authoritative definition of the status of its subject and the secure repository of the work of a generation. Much of the present volume qualifies in this sense, though not perhaps quite all of it.

Experimental psychologists will find this text invaluable. The chapters on the senses, in particular, provide an up-to-date account of modern work of the greatest interest and value. At the same time, it should be said that—with some notable exceptions—there is little direct reference to the findings of psychological experiment. It is understood, however, that contributions of more immediate relevance to psychology are to be included in the later volumes.

It is the evident hope of its sponsors that this new venture will find a place in the long and praiseworthy tradition that extends from Haller's *Elementa* to the formidable multi-volume German handbooks of the nineteenth and early twentieth centuries. There is, however, one unfortunate difference. The prevailing division of the modern world has led to practically no account being taken here of contributions to neurophysiology from the other side of the Iron Curtain. Unless the omission is significantly remedied in later volumes, this *Handbook* will inevitably come to be regarded over large sections of the globe as a Handbook of Western Physiology. To those who believe in the unity of science this might seem a pity.

This volume is a major contribution to science and scholarship. All who have been concerned in its production are to be congratulated on a remarkable achievement.

O. L. Zangwill.

Research Techniques in Human Engineering. By Alphonse Chapanis. Baltimore. Johns Hopkins Press. 1959. Pp. vii + 316. $6.

This is a book whose purpose is avowedly to present methods rather than results. As such it fills a definite need since much of the application of psychology to practical problems requires an experiment for each problem rather than the use of knowledge picked out of a reference book. In many cases, therefore, there is a need for people working in industry to acquire the research approach, and to learn to apply experimental methods to their own problems. Indeed there are too few psychologists to carry out all the experimental work which needs doing on simple *ad hoc* problems; even if this were not a misuse of an academic training in the subject. One may therefore hope that Professor Chapanis' book may help to disseminate the experimental approach to behaviour.

The various sections of the book deal with methods of direct observation (activity sampling and other work study techniques), with methods of studying accidents (such as the critical incident technique) with statistics, with experimental design, with general research strategy, with the psycho-physical methods, and with articulation testing techniques. As this list shows, the treatment is fairly exhaustive, and the presentation of each section is clear and accurate. Each general point is illustrated by a concrete example of an investigation to which it was relevant. There should be no difficulty for engineers and other non-biologists in absorbing the information Professor Chapanis provides.

The subject matter would doubtless have been rather different in a book produced on this side of the Atlantic. In England applied psychology tends to be associated with applied physiology more closely than it is with work study; so that one would find it surprising to meet a complete list of the Therblig categories in a book which contains no information on the measurement of energy consumption at all. Equally, the allocation of a complete chapter to articulation testing does not reflect the relative importance of that type of problem in British application. Some of us may also regret the subject matter of the statistical chapters, since they provide a clear exposition solely of methods based on the analysis of variance. Recent criticism of such methods has been so severe that many investigators are abandoning the use of Latin Square designs, and are using non-parametric methods wherever possible. Even if this is felt to be an extreme course, it would surely be worth giving such simple methods as the Sign Test in a book which goes so far as to consider the extraction of fourth order interactions.

But it is likely that, quite apart from cultural differences, any individual applied psychologist would produce a different list of topics to be included in a book with this general aim. Professor Chapanis is to be congratulated on his success in producing a compendium of research techniques which is simple enough for the complete novice, and yet sufficiently detailed to be of value for reference by the professional.

D. E. BROADBENT.

What is Cybernetics? By G. T. Gilbaud. (Translated by Valerie Mackay from *La Cybernetique*. Paris, Presses Universitaires de France, 1954.) London. Heinemann: Contemporary Science Series. 1959. Pp. viii + 126. 10s. 6d.

A non-technical introduction to cybernetics has long been needed, and this reasonably priced volume should go far to fill the gap.

Most of the book is devoted to a general discussion of information, its measurement, and the laws of communication, with several useful examples worked out numerically. Feedback and network theory are dealt with more shortly, and no account is given of the actual properties of closed-loop systems. In particular, no mention is made of the important effects of delayed feedback. The theory of games is very briefly mentioned in the final chapter.

The selection and treatment of topics, and the authorities cited, will seem refreshingly different to the Anglo-Saxon reader. Most of the important cybernetic ideas seem to have appeared quite early in French mathematical and philosophical writing, to be embodied later (often unwittingly) in English and American engineering. To judge from this book the French still prefer a broad intellectual synthesis rather than a concrete problem-solving approach.

The translation is excellent, though certain technical terms have been transliterated rather than translated. For instance, Figure 2 (p. 13) would normally be called a "switch" rather than an "interrupter."

E. R. F. W. CROSSMAN.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XII 1960 Part 4

CHOICE REACTION TIMES FOR SKILLED RESPONSES*

BY

G. H. MOWBRAY

From the Johns Hopkins University Applied Physics Laboratory, Silver Spring, Maryland

Two kinds of choice reaction time experiments are reported, both of which make use of a highly overlearned sensori-motor response. When a response is required for each stimulus presented, no increase in reaction time occurs as a function of the number of alternative stimuli available. It is proposed that the increase in choice reaction times commonly thought to accompany an increase in the number of alternative choices provided reflects the unpractised state of the responder. When a response is required for only one out of *n* possible stimuli, a slight but consistent increase in reaction time takes place with an increase in the number of alternatives. An analogy is drawn between the second experiment and a vigilance task and an expectancy hypothesis is invoked to explain the results.

Introduction

It has long been accepted as fact that choice or disjunctive reaction times increase with the number of alternative responses required. The generality of this dictum apparently has been accepted without question since the early experiments of Merkel (Merkel, 1885; Woodworth, 1938). Further impetus to its acceptance has been provided in recent years by those interested in applying communication theory techniques to the analysis of human behaviour (Hick, 1952; Crossman, 1953; Hyman, 1953). However, as data have accumulated from diverse investigators using different techniques of stimulus presentation and response, it has become increasingly clear that qualifications to the basic generalization are needed. As an example, Pierce and Karlin (1957) have shown that reading rates do not increase when the size of the vocabulary used is increased, and most recently Leonard (1959), in a cleverly conceived experiment, has produced data on tactual choice reactions that show no increase in choice reaction time as the number of alternatives is increased from two to eight. Previous to Leonard's finding, Mowbray and Rhoades (1959) had demonstrated that the difference in reaction times between a two-choice and a four-choice finger response could be reduced to zero by intensive practice. Finally, Saslow (personal communication), using recorded speech syllables as stimuli and repetition of the stimulus as the response, has obtained reaction times that do not change as a function of the number of alternative responses required.

One question that arises from these results might be stated in the following way. Is the reported finding that choice reaction times increase with increasing numbers of alternatives a reflection of the unpractised state of the responders when a complex neuro-muscular response is required? It is certainly true that those

* This report was prepared under Contract NOrd 7386 between the Bureau of Naval Weapons, Department of the Navy, and The Johns Hopkins University.

experiments that show the greatest effect due to increased numbers of alternatives also utilized the most unnatural or unfamiliar stimulus-response combinations (cf. Merkel, 1885; Hick, 1952; Crossman, 1953; Hyman, 1953). On the other hand, in those cases where small increases or no increases at all were encountered, the subjects were either highly practised or else familiar stimulus-response combinations were employed (cf. Crossman, 1955; Pierce and Karlin, 1957; Mowbray and Rhoades, 1959; Leonard, 1959).

A shortcoming common to nearly all of the researches carried out to date concerns the small number of subjects tested. Understandably, it is tedious and sometimes prohibitively uneconomical to use large numbers of observers where a vast amount of practice is required or where many experimental conditions are used. However, it is dangerous to generalize from a limited sample.

Some of the above considerations were uppermost in mind when the experiments to be described were designed. To avoid the tedium of long practice sessions, a task was sought that required a fairly complex neuro-muscular response to a stimulus that was compatible with the response. The task finally chosen as the one most nearly meeting all requirements was the vocal identification of visually presented Arabic numerals. Undeniably that is a perceptual-motor skill that gets thoroughly over-learned by most literate people. By using this type of task it was hoped to achieve two things. First, by taking advantage of an already highly over-learned skill, it was reasoned that a more definitive test could be made of the hypothesis that choice reaction time is an increasing function of the number of alternative responses required. Second, the savings in time afforded by the use of an over-practised perceptual-motor response would permit the use of a larger number of observers than has been customary to date.

It should be noted at this point that Alluisi, Muller and Fitts (1957) report data on the vocal naming of visual numerals in a forced-paced serial task. Interestingly enough they found no differences between relative rates of information transmission with changes in the number of possible alternative stimuli over a range of 1 to 3 bits/stimulus. They were, however, not concerned with reaction times to individual stimuli nor did they even record such data.

Experiment I

Apparatus and procedure

The apparatus consisted of a teletype, punched-tape transmitter and a bank of coding relays that supplied the anode voltage to a Burroughs read-out tube. These tubes, called Nixies, contain tiny wire filaments in the shape of the numerals 0–9, and when the circuit through any one of these filaments is completed, the appropriate number glows brightly enough to be seen easily in a normally lighted room. They are a gas-filled, cold cathode tube whose light-up time is virtually instantaneous. They extinguish equally rapidly. The numeral they produce is 0·6 in. in height with a stroke width-to-height ratio of one to eight.

Standard electronic timing circuits generated a pulse that triggered the tape transmitter every 2·5 sec. so that a new stimulus appeared that often. The transmitter triggering pulse also operated the clutch of a Standard Timer Model S-1. The clutch of the clock timer was de-energized and the clock stopped when the subject responded by naming the lighted numeral. This was accomplished by an electronic voice key operated from a throat microphone worn by the subject. All voice reaction times were measured to the nearest one-hundredth of a second. Delays caused by relay inertia were measured and found to be 65 millisec. on the average. Consequently, all reaction times reported were corrected by that amount.

For this experiment five different degrees of choice were provided—2, 4, 6, 8, or 10 choices—and any one subject was assigned randomly, with certain restrictions, to only one of the five possible conditions. Thus a subject who was assigned to the two-choice

condition saw, and of course, responded to one of two possible numerals. Reaction was the simple vocal naming of the numeral that appeared on the read-out tube. A total of 150 subjects participated in the experiment with the only restriction on the randomness of assignment to conditions being that an equal number of males and females be represented in each choice level. Therefore, there were fifteen males and fifteen females in each of the five choice conditions under test. The subjects, with the exception of a few laboratory technicians, clerical assistants and college faculty members, were college students. Their median age was 19·5 years with a range of 17–48 years. Most had never previously served in any type of psychological experiment. Table I shows a distribution of the ages of all of the subjects for each of the five experimental conditions.

TABLE I

DISTRIBUTION OF THE AGES OF THE 150 SUBJECTS FROM EXPERIMENT I FOR ALL EXPERIMENTAL CONDITIONS

Age

Experimental conditions	17–20	21–24	25–28	29–32	33–36	37–40	41–44	45–48
2 choices ..	19	7	2	1	1	—	—	—
4 choices ..	14	11	2	1	1	—	—	1
6 choices ..	17	9	2	—	—	2	—	—
8 choices ..	20	7	1	2	—	—	—	—
10 choices ..	18	9	3	—	—	—	—	—

The subjects were seated facing the read-out tube and at a distance of 7–8 ft. with the tube at approximately eye level. Before each experimental trial commenced, the subject was told how many numbers he would see and what those numbers were. Stress was placed on the fact that those were the only numbers he would see, and if there appeared any doubt or hesitation it was made certain that the subject understood the task and, more important, that he knew which numbers he was going to be shown. Each experimental run was composed of 60 repetitions of each of the numerals included in a particular condition, thus insuring equal probability of appearance for all numerals. Therefore, subjects assigned to the two-choice condition provided 120 responses whereas those assigned to the ten-choice condition provided 600. The order in which the numerals appeared was in all cases random.

In analyzing the results only the reaction times to the appearance of the numeral 8 were used since this was the only number common to all conditions. The 8 was chosen originally because, beginning with a vowel sound as it does, it provided more immediate acoustic power for operation of the throat microphone than most of the other numerals. Finally, only the responses to the last 50 key stimuli were used—not counting errors. This procedure allowed for a brief but sufficient familiarization and warm-up period.

TABLE II

MEAN CHOICE REACTION TIMES AND STANDARD DEVIATIONS FOR 150 SUBJECTS AS A FUNCTION OF THE NUMBER OF CHOICES INVOLVED. TIMES ARE IN SECONDS

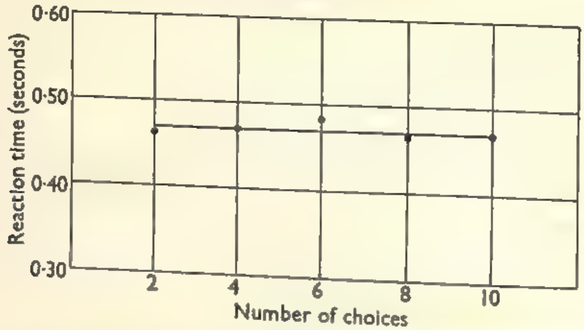
Number of choices	2	4	6	8	10
Mean reaction time ..	0·461	0·469	0·481	0·463	0·464
Standard deviation ..	0·075	0·060	0·045	0·078	0·065

Results

The mean choice reaction times and standard deviations obtained for the various degrees of choice are given in Table II. The same data are plotted in Figure 1 where the least squares fit to the data points is shown. As can be seen, the slope

of the least squares fit line is zero. That demonstrates, in effect, that choice reaction time is a constant for an over-learned response where as many as ten choices are involved.

FIGURE 1



Mean choice reaction time as a function of the number of choices available to the responder. Each data point is the mean of 50 reaction times from 30 different subjects.

The reaction times of all subjects were submitted to further analysis to determine whether the vocal naming response utilized in this experiment was indeed an over-learned one. To accomplish this, the mean times for the first 25 scored reactions were compared with the mean times for the last 25 reactions for all subjects and all degrees of choice. The mean differences so obtained were then tested for significance using the method recommended for correlated means (McNemar, 1949). None of these differences in any way approached significance. Thus it can be concluded that the response being measured was over-learned to the point that no further improvement took place throughout the course of the experiment.

Finally, a comparison was made between the reaction times of the male and female subjects. It was found that the females on the average reacted more slowly than the males by a nearly constant amount. The average difference of approximately 50 millisec. was significant at beyond the 0·001 level of confidence.

Experiment II

The experiment described above has dealt with but one of the two classical types of choice or disjunctive reaction. The other type requires the subject to react to only one stimulus of two or more that might be presented to him. Thus, in this latter case, the subject knows what the reaction will be; however, he does not know when it will be required. A second experiment of this latter sort was designed and performed. The apparatus and the experimental material were the same as in Experiment I. The subjects were 74 of the 150 who had served originally, but they were assigned to a different experimental condition than the one they had served in previously. Originally, it had been planned to use 75 subjects but due to a scheduling error only 74 were run, so that there were finally 15 subjects in each of 4 of the experimental conditions and 14 subjects in the fifth condition. There were 37 males and 37 females divided nearly equally between the 5 experimental conditions.

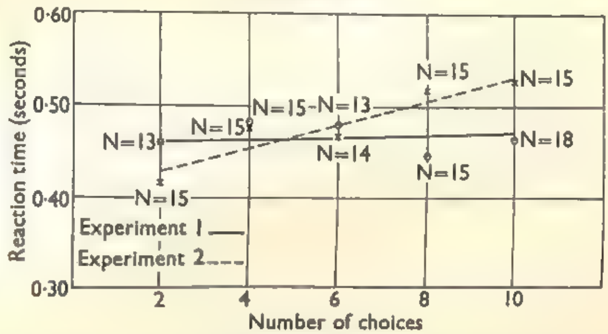
The procedure was identical with that described for Experiment I with one important exception. In this second experiment the subjects were required to respond only to the appearance of the figure 8. Thus, although they saw all of the numbers that were used for a given choice condition they could ignore, in so far as responding was concerned, everything except the appearance of the 8. This stipulation provided for the subject one complication that was not present in the first experiment. It caused an uncertainty about the time that a response would be required. In Experiment I, a response was required every 2·5 sec., but in this latter case responses were randomly distributed with a mean response interval that increased as a function of the number of choices.

Results

The mean reaction times for all subjects who participated in Experiment II are plotted in Figure 2 as a function of the number of choices involved. They are the points connected by a dotted line. Also plotted, as a solid line, are the mean reaction times of the same subjects from Experiment I. The *N*'s for the data from Experiment I differ somewhat from point to point because the subjects for Experiment II were selected at random from among those who had participated in Experiment I. They were not selected on the basis of what condition they had seen previously. It was purely accidental that the *N*'s turned out to be so nearly similar.

It can be seen in Figure 2 that the reaction times in general show an increase as the number of choices increases. Analysis of the data using the F-ratio test for linearity of regression yielded a ratio that could have been obtained less than one time in a thousand on a chance basis. Thus it can be concluded that the means from Experiment II differ as a function of the number of choices involved. The mean reaction times that these subjects provided from Experiment I (solid line in Figure 2) were submitted to a similar analysis and they were found not to differ as a function of the number of choices.

FIGURE 2



Mean choice reaction times from Experiment II compared with those from Experiment I. N represents the number of subjects contributing reaction times to a given choice condition.

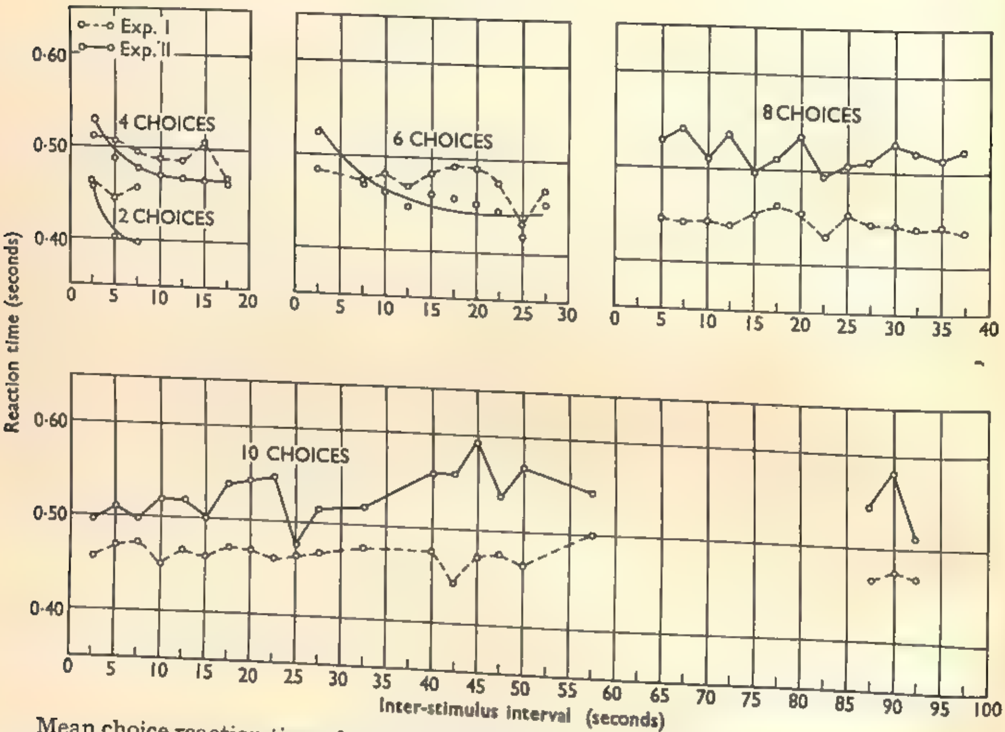
Once again the results from the male and female subjects were separated. As previously, the reaction times of the females were significantly slower than those of the males. It was also noted that the females showed a tendency to be more affected by the number of choices than did the males. As a matter of fact, a statistical check of the results from the male subjects for departure from zero slope yielded doubtful significance ($P = 0{\cdot}05$). A similar test on the data provided by the females resulted in a P of 0·001.

The results outlined so far from both experiments indicate that requiring a subject to respond to every one of several alternative stimuli is somehow different from requiring him to respond to only one stimulus out of several possible alternatives. Some further analyses were carried out on the data from both Experiment I and Experiment II in an attempt to gain some insight into the reasons for this difference.

In physical construction, the two experiments were identical, but in operation, it will be remembered, there was one major point of dissimilarity. Experiment I called for a response every 2·5 sec., whereas in Experiment II a response was required only when the key figure 8 appeared. Thus in the latter case, responses were generally random with the time interval between responses increasing on the average as the number of choices increased.

In order to determine whether the increasing inter-stimulus interval had contributed to the differences observed, an analysis of the reaction times to the key stimulus as a function of this interval was made. The reaction times from all subjects in Experiment II were used as well as the reaction times from Experiment I of those subjects who also served in Experiment II. The results from both experiments are separately plotted in Figure 3.

FIGURE 3



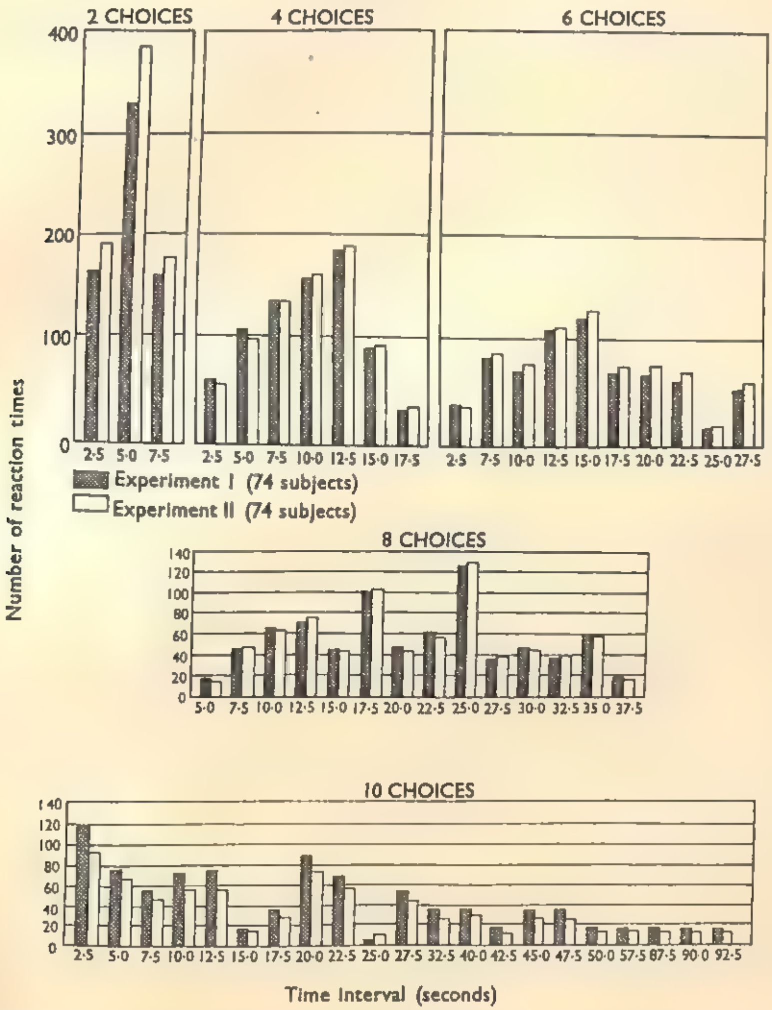
Mean choice reaction times from Experiment I and Experiment II plotted as a function of the inter-stimulus interval for all conditions of choice.

It is important to notice first that the data from Experiment II for 2, 4, and 6 choices show an unmistakable trend downward as the intervals between stimuli increase. This trend is not evident for 8 and 10 choices nor, apparently, for the data from Experiment I. Another thing to notice, and one that shows up grossly in Figure 2, is the fact that the reaction times for 2 choices and 4 choices are generally higher overall for Experiment I, whereas 8 and 10 choices yield higher reaction times for Experiment II. The six-choice condition seems to be the transition point although the bulk of the scores from Experiment I are higher.

It is clear that the actual distributions of intervals for any given choice condition are important in interpreting the data from Figure 3. These distributions for the two experiments are provided in Figure 4. Unfortunately this type of analysis was not anticipated when the experiments were designed, and there are thus too few data at certain of the intervals to be anything more than suggestive. The distributions of the intervals for the three lowest choice conditions are, however, fairly regular and have enough reaction times at most intervals to be dependable. For the two-choice condition, where the data for each interval used are abundant, the discrepancy

between the reaction times from the two experiments is marked. The same sort of discrepancy appears in the four-choice condition with the exception of the 17·5 sec. interval which is somewhat unreliable due to the fact that it contains only two reaction times for each subject. For the six-choice condition the discrepancy, while not marked, is nonetheless evident. A discussion of what is judged to be the significance of these points will be postponed for the moment.

FIGURE 4



The distributions of inter-stimulus intervals employed for all choice conditions in both experiments.

One further analysis of the data from Experiment II was made. It consisted of a comparison between the mean reaction times for the first 25 reactions and the last 25 reactions for each choice condition. The same comparison was made with the data from Experiment I and, it will be remembered, no significant differences were found. In this latter case, however, the reaction times from the second half of Experiment II were significantly higher than those from the first half ($P < 0{\cdot}01$).

The differences between the least squares fits to the data points were approximately 12 millisec. for 2 choices and 28 millisec. for 10 choices.

Discussion

The results of Experiment I indicate very convincingly that choice reaction times for the type of stimulus-response complex used do not increase as a function of the number of alternatives available to the responder. To put the matter in a more positive light, the results support the hypothesis that choice reaction times for practised responses are constant. Interestingly enough, Saslow has recently obtained similar results using a task that required the vocal naming of orally presented syllables. Even more recently Seibel (1959) has reported that extended practice of finger responses to the appearance of light patterns resulted in no consistent relationship between reaction time and the number of alternative patterns for as many as 30 alternative patterns.

In view of the results available from many sources now it is evident that the generalization regarding the increase of choice reaction times as a function of the number of alternative responses required has to be abandoned. Apparently it arose originally as a reflection of the unpractised state of the responders and has been perpetuated in more recent times because it fitted the hypothesis of man as an information processing machine.

The mean, overall reaction times from Experiment I are somewhat higher for the lower choices than those usually encountered in similar research suggesting the possibility that the results are due to slow responding on the part of the subjects in the lower choice conditions. Hick (1952), for example, reported his reaction time to a two-choice task to be between 250 and 260 millisec. and Merkel (quoted by Woodworth, 1938, p. 333) gives 316 millisec. as the time for a two-choice reaction. However, at the other end of the scale, Hick's ten-choice time was over 525 millisec. and Merkel's was 622 millisec. This is some 100–150 millisec. slower than the times reported here. Since voice reaction times are considered to be much slower than hand or finger reaction times (Woodworth, 1938, p. 329), the assertion that previous findings are due in large part to improperly practised subjects is further strengthened.

On the assumption that the fastest responders would be less likely to be influenced by any factors inherent in the test situation that tended to cause a slowing down of response to the lower conditions of choice, the scores from the three fastest responders for each choice condition were examined separately. The means for the two-choice condition and the ten-choice condition were found to differ by only 13 millisec. (365 millisec. as against 378 millisec.). While not entirely conclusive, perhaps, this result at least does nothing to negate the central thesis.

Fortunately, or unfortunately as the case may be, man is adaptable. Faced with an unfamiliar and complex sensory motor choice task, he attacks it as efficiently as he can, perhaps by some sort of progressive classification or self-replication process as suggested by Hick (1952). However, with repetition he improves until mastery of the task is complete. At this point he appears to have acquired a reflex. There is no longer an increasing relationship between the reaction time and the number of alternatives; that is to say, choice reaction time is a constant.

The complication provided by the results from Experiment II does not enhance the strength of the position adopted above. Yet, it must be remembered that from the subject's point of view these were two extremely different situations. The question that needs to be resolved, quite obviously, is what factors were operating in the second experiment that could have caused such differences as were found? To this writer, the most intriguing difference is the one that shows up for the three

lower conditions of choice in the time interval analysis (see Figure 3). When the subject had to respond only to the appearance of the key figure, as in Experiment II, his reaction times show a steady and consistent decline as the time interval between key stimuli increased. This suggests that the subject was operating consciously or unconsciously on an expectancy basis. An argument of this type would predict that the subject would not expect a key stimulus immediately following another key stimulus and thus his expectancy would be low, resulting in a higher reaction time. As the time since the last key stimulus increased, expectancy, based no doubt on experience with the task, would increase also, thus providing lowered reaction times. A look at the distributions of intervals encountered, shown in Figure 4, indicates that the subject would be thoroughly justified in adopting such a mode of behaviour.

Deese (1955) has proposed an expectancy hypothesis in an attempt to account for some of the facts of vigilance, and it seems worthwhile to consider this in relation to the results of the second experiment since there are some similarities between that experiment and a vigilance situation. Both tasks require that the subject be on the alert for a signal whose time of appearance is unpredictable for the most part. Furthermore, rate of signal appearance is a common variable in vigilance research that also appears in the reaction time study.

Two predictions about watch-keeping behaviour that stem from an expectancy hypothesis are the following: (*a*) the probability of detection of a signal should be low immediately following a signal and should improve as the inter-signal interval increases; (*b*) the overall probability of detecting a signal should be high for high signal rates and low for low signal rates. To further the proposed analogy between a vigilance task and the task provided by Experiment II it is only necessary to assume that probability of detection and response latency would behave similarly under similar conditions and, further, that the effects observed over long periods of time would also show up with the more abbreviated time intervals encountered in these experiments. The first of these assumptions has some experimental support. Jenkins (1954) provided his subjects in a vigilance experiment with a regular but unscored signal to which they had to respond by pressing a key. He found that latency of response was clearly associated with probability of detection. When signal rates were high and probability of detection was high, response latencies were low, and conversely, when signal rates were low and probability of detection was low, response latencies were high. The second assumption is merely a logical extension of Deese's expectancy hypothesis to include shorter time intervals.

Viewed in the light of the above discussion, the results from Experiment II are in fair accord with the predictions generated from an expectancy hypothesis. Certainly the three lowest conditions of choice show an unmistakable lowering of the reaction time as the inter-stimulus interval increases. Why the eight-choice and ten-choice conditions do not show this same effect is not known. It is perhaps reasonable to attribute the cause to the irregular time interval distributions at those two points and to the scarcity of data at many of the intervals. Further research is needed to resolve this question.

The data from this experiment conform also to the second prediction of the expectancy hypothesis in that the reaction times are low for high signal rates and high for low signal rates. In one other respect, too, these results are similar to those commonly encountered in vigilance research. Specifically, there is evidence of deterioration of performance with respect to the duration of the task. The reaction times for the second half of the experiment were significantly slower than for the first half. Furthermore, the low signal rates appear to be more affected than high

signal rates although it is not meant to imply that the small differences encountered here are anything more than suggestive.

In summary, then, it is suggested that the differences observed between these two experiments are a direct outcome of requiring a response to every stimulus in one case and to only one of *n* stimuli in the other. In the first instance a constant inter-stimulus interval and a constant stimulus rate resulted in constant overall latencies. Apparently the number of alternative responses available to the responder had no direct bearing on response latency. In the second instance there is strong, if not conclusive, evidence that the variables of inter-stimulus interval and stimulus rate operated to produce latencies that decreased as the interval increased and increased as the stimulus rate decreased. An expectancy hypothesis as outlined by Deese to interpret some of the facts of vigilance research is invoked and found to be reasonably adequate. However, any more positive conclusions must await future research using a variety of distributions of inter-stimulus intervals and stimulus rates.

References

Alluisi, E. A., Muller, P. F., Jr., and Fitts, P. M. (1957). An information analysis of verbal and motor responses in a forced-paced serial task. *J. exp. Psychol.*, **53**, 153–58.

Crossman, E. R. F. W. (1953). Entropy and choice time: The effect of frequency unbalance on choice response. *Quart. J. exp. Psychol.*, **5**, 41–51.

Crossman, E. R. F. W. (1955). The information capacity of the human operator in symbolic and non-symbolic control processes. In J. Draper (Ed.). *The Application of Information Theory to Human Operator Problems.* W. R. (D.) Report No. 2/56, Ministry of Supply.

Deese, James (1955). Some problems in the theory of vigilance. *Psychol. Rev.*, **62**, 359–68.

Hick, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11–26.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, **45**, 188–96.

Jenkins, H. M. (1954). Unpublished paper read in a Symposium on Vigilance before the American Psychological Association in New York City, U.S.A.

Leonard, J. Alfred (1959). Tactual choice reactions: I. *Quart. J. exp. Psychol.*, **11**, 76–83.

McNemar, Q. (1949). *Psychological Statistics.* New York.

Merkel, J. (1885). Die zeitlichen Verhältnisse der Willensthätigkeit. *Philos. Stud.*, **2**, 73–127.

Mowbray, G. H., and Rhoades, M. V. (1959). On the reduction of choice reaction times with practice. *Quart. J. exp. Psychol.*, **11**, 16–23.

Pierce, J. R., and Karlin, J. E. (1957). Reading rates and the information rate of a human channel. *Bell System Tech. J.*, **36**, 497–516.

Seibel, R. (1959). Discrimination reaction time as a function of the number of alternatives and of the particular stimulus-response patterns. *Amer. Psychol.*, **14**, 396.

Woodworth, R. S. (1938). *Experimental Psychology.* New York.

STUDIES IN OLFACTORY ACUITY

III. Relative Detectability of *n*-Aliphatic Acetates by the Rat

BY

D. G. MOULTON*

From the Department of Psychology, Stanford University, Stanford, California

The ability of black and albino rats to detect the first seven *n*-aliphatic acetates in the vapour phase has been investigated by use of an odour choice box. Both the median threshold and the slope of the probit regression line (which describes the relation between concentration and response) tend to decrease logarithmically in value as the series is ascended. The position of methyl acetate in these relations appears anomalous. When expressed as thermodynamic activities median thresholds for acetates of intermediate chain length are about equally stimulating, whilst short chain alcohols excite at decreasing activities with increasing chain length. The close similarity of these findings to those reported in the previous study on *n*-aliphatic alcohols supports the view that such relations may be of more general occurrence in olfaction in the rat.

The stimulus-response curve for each acetate studied shows one or more reversals in slope. It is suggested that certain of these are due to the presence of two or more olfactory receptor types one or more of which attain maximum response at lower concentrations of a given acetate than do the remainder. The asymptotes of the curves are reached in about 2–3 log units of concentration.

Introduction

There is growing evidence that in smell as in taste (Beidler, 1954; Dethier, 1956) the events leading to the excitation of the chemoreceptors are more probably biophysical than biochemical in nature. In particular, recent studies of olfaction in homologous series, in which behavioural and electrophysiological methods have been used, have emphasized the importance of properties such as solubility which alter progressively in value within a series (e.g. Dethier and Yost, 1952; Ottoson, 1958). Thus in a previous report, describing the responses of rats to alcohols, it was found that detectability at median threshold rises logarithmically with increasing chain length (Moulton and Eayrs, 1960). It can also be argued that the tendency noted by these and other workers for intermediate members of certain homologous series to be equally effective odorants at about equal thermodynamic activities is further evidence that the dominant process in the overall effect is physical.

However, since many physical properties vary in a similar manner as chain length is expanded, consideration of their relations to odour intensity in a single series will throw limited light on their relative significance for olfaction. As the important studies of Dethier and Chadwick on tarsal chemoreception in the blowfly (see Dethier, 1956) have clearly illustrated, information of this kind is more likely to be derived by an analysis of such relations for a number of series in which different substitutions in the same carbon skeleton do not alter all properties in the same way. Evidence which would permit such an analysis is scarce and relates to species other than the rat. It is therefore one aim of this study to provide data on the detectability by the rat of a further series of homologous odorants.

A second aim concerns the finding of the previous report that in alcohols the slopes of the probit regression lines (which describe relations between concentration and response) decrease logarithmically in value as the series is ascended. Since this

* Present address: Division of Physiology, Florida State University, Tallahassee.

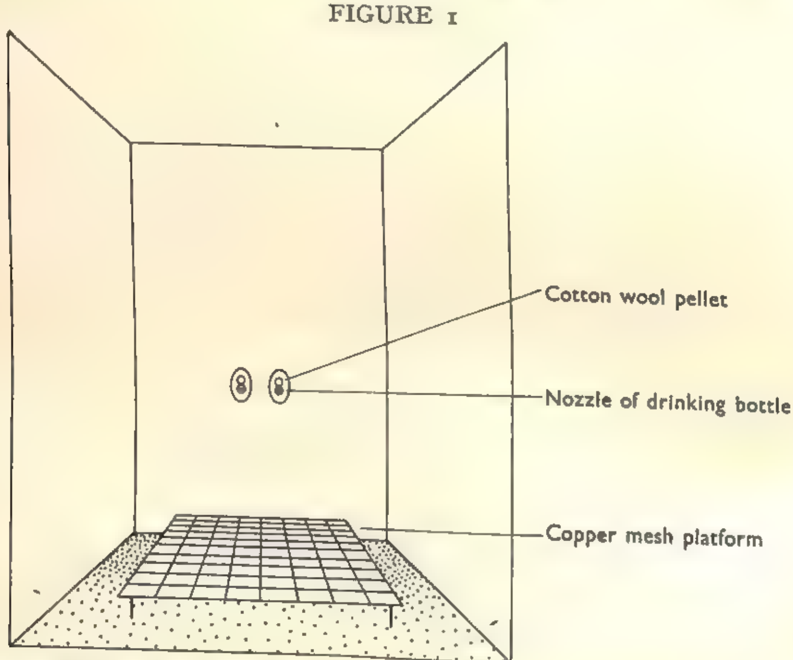
implies that there is a progressive change in the pattern of detectability as the concentrations are increased and therefore may explain the apparently conflicting claims in the literature about the pattern of detectability in homologous series, it is important to know whether this trend is unique to the alcohols or represents a more general phenomenon. The following study has therefore been designed to provide data over a wide range of the stimulus-response relations in the first seven n-aliphatic acetates.

Materials and Methods

Apparatus

All rats were initially trained in the circular choice apparatus described by Eayrs and Moulton (1960), but in an attempt to reduce the time required to test rats a modified form of the rectangular choice box (also described by Eayrs and Moulton) was used in definitive trials (Fig. 1.) This version has three hardboard walls and a plastic floor

FIGURE 1



Odour choice box (for description, see text).

supporting a copper mesh platform. Both floor and platform can be removed for cleaning. Two bottles are hung behind the wall facing the open side of the box so that their spouts protrude through two apertures above the platform. These spouts are identical to those used in the circular choice apparatus, each bearing a holder into which can be fitted a detachable nylon capsule. A cotton wool pellet soaked with either the odorant or the control solution is placed in each capsule. Both tubes are connected by wires and clips to an "on-off-on" switch and hence to one pole of a shocking device whilst the other pole is connected to the platform on which the rat stands. In this way either spout can be charged so that a rat drinking from it receives a shock. The position of the negative capsule, containing the odorous pellet is varied according to a randomly determined sequence so that a given dilution of a test compound appears as often on the right-hand position as on the left. Other details of procedure and scoring of trials were as described by Eayrs and Moulton (1960).

Animals

In the main experiment 42 black and albino male rats, taken from a stock initially derived by crossing black and albino strains of the Norway rat (*Rattus norvegicus*, Berkenhout), were used. All rats had been trained to detect members in the alcohol series and

18 of them were subjects in the previous study. At the beginning of the experiment they varied in age from about 11–13 months.

In order to determine the ease with which experimentally naïve rats could be trained, a group of 5 young albino and black rats were also tested in the odour choice box after they had been given two nights training on a modified circular choice apparatus. The test odorant was a concentrated solution of butyl bromide and the rats were not used in investigating the detectability of *n*-aliphatic acetates.

Odorants

All reagents used were obtained from the Eastman Kodak Company. The acetates were diluted with propylene glycol by successive factors of two, essentially according to the method described by Beck, Kruger and Calabresi (1954), to form a geometric series of dilutions.

Experimental design

In an attempt to minimize the influence of day to day fluctuations in performance the series of dilutions chosen for each acetate was presented to the rat in accordance with a balanced incomplete block design. Thus nine dilutions were tested in a series of twelve blocks, each of which consisted of three dilutions presented in a randomly determined manner. This allowed four replications of each dilution to be made. 42 rats were tested on each of the acetates which were presented in the order: amyl, ethyl, butyl, methyl, propyl, heptyl, hexyl. The appearance of anomalies in the stimulus-response curves of these acetates suggested that additional data would be of value. Replications for three acetates were therefore made using identical procedures and the same number of rats, but employing fresh sets of dilutions.

Treatment of data

The results were subjected to probit analysis (Finney, 1952) and values for median threshold response were determined as molar concentrations in solution. The procedure used for estimating the concentration of the alcohol vapour and the thermodynamic activities corresponding to these values were as described by Moulton and Eayrs (1960).

Cheesman and Mayne (1953) have pointed out that the application of probit analysis to data such as the present, results in anomalies which have the effect of setting fiducial limits at artificially low values. For this reason fiducial limits have not been used in assessing the results.

Results

Technique

Rats which had previously been trained and tested in the circular choice apparatus rapidly learned the discrimination when transferred to the odour choice box, whilst the group of 5 experimentally naïve rats reached the criterion of 10 consecutive correct responses in 38 trials or less. Fully trained rats could be tested much more rapidly in the odour choice box than in the circular choice apparatus.

TABLE I

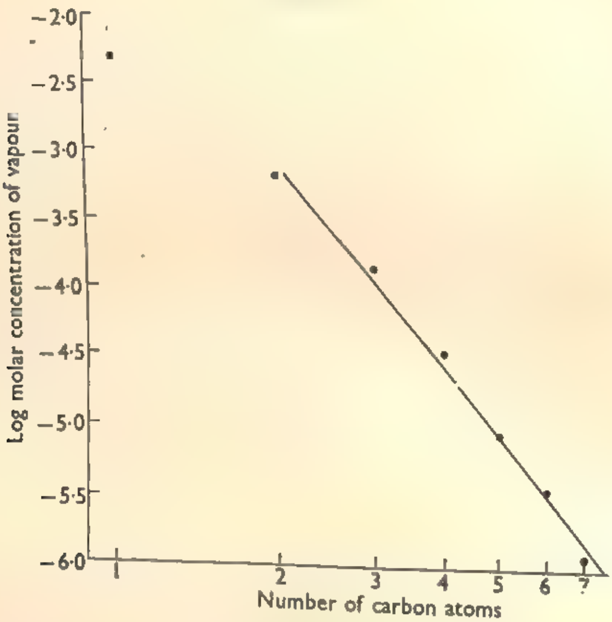
Detectability of *n*-Aliphatic Acetates in the Vapour Phase by 42 Rats. Each Rat Received 36 Trials on Each Acetate

Acetate	*Slope of regression line*	*Log molar concentration of solution*	*Log molar concentration of vapour*	*Log activity*
Methyl ..	1·037	0·7720	−2·3267	−0·3676
Ethyl ..	1·396	0·2900	−3·1891	−0·8348
Propyl ..	0·776	0·0030	−3·8536	−1·1206
Butyl ..	1·025	−0·1410	−4·4746	−1·2634
Amyl ..	0·534	−0·1850	−5·0429	−1·3047
Hexyl ..	0·601	−0·1815	−5·4437	−1·2970
Heptyl ..	0·390	−0·2337	−5·8623	−1·2459

Relative detectability of n-*aliphatic acetates*

It is apparent from the results given in Column 4 of Table I that the median threshold decreases logarithmically as chain length is expanded and if the value for methyl acetate is excluded there is clearly a very high degree of correlation. This fact is emphasized in Figure 2 in which molar concentrations at median threshold are plotted against the log of the number of carbon atoms in the alkyl group of the molecule. The values of the slopes of the probit regression lines given in Table I also show a linear relation to chain length and although the correlation is less striking the tendency for slopes to decrease logarithmically in value as the series is ascended is significant at the two per cent. level. This relation implies that the detectability pattern of the acetates varies with concentration.

FIGURE 2



Relation of median threshold concentration of an homologous series of acetates, expressed as log molar concentration of vapour, to the number of carbon atoms in the alkyl group of the molecule.

Since there are grounds for expecting the first member of a series to show anomalous behaviour the value for methyl acetate has been excluded from the computation of the regression line shown in Figure 2 (cf. Moulton and Eayrs, 1960).

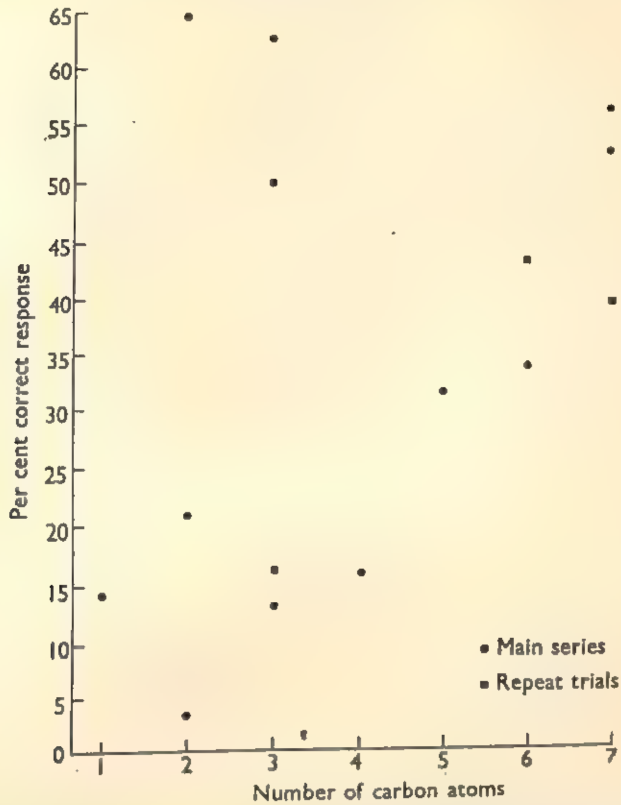
Anomalies in the stimulus-response curves for the acetates

In the stimulus-response curves of the acetates studied there is a series of reversals in slope. An example of this is seen in the data for amyl acetate plotted in Figure 5, which is more conveniently placed in the discussion. If the percentage correct score corresponding to the lowest point in the trough of each reversal is plotted against chain length as is done in Figure 3, the distribution of these reversals in relation to their position on each curve becomes apparent. (A similar pattern results when the peak instead of the trough of a reversal is used as a criterion.) From this it appears that in members above ethyl acetate only one reversal is consistently present on successive replications and this shows a tendency to occur at progressively higher

response levels as the series is ascended. In fact there is a significant correlation between the position of these reversals on the curve and chain length ($p < 0{\cdot}001$).

The asymptotes of the curves are reached in about 2–3 log units of molar concentration.

FIGURE 3



The position of reversals of slope on the stimulus-response curves of an homologous series of acetates. The percentage correct score corresponding to the trough of each reversal is plotted as a function of the number of carbon atoms in the alkyl group of the molecule.

Discussion

Technique

It has been shown that the circular choice apparatus can be used effectively to train and test rats in an odour choice situation (Eayrs and Moulton, 1960). However, when an extended series of trials is undertaken the routine cleaning of the apparatus and the test procedure becomes laborious and time consuming. The present results indicate that these difficulties can be greatly reduced and the advantages of automatic training largely retained if preliminary training on the circular choice apparatus is followed by use of the modified odour choice box for routine testing. Thus the testing procedure is much simplified while the discrimination is learned rapidly and without the appearance of behavioural disturbances such as occurred during trials with the rectangular choice box (Eayrs and Moulton, 1960). Further, the box is easily incorporated into an olfactometer utilizing air dilution of the odorant (Moulton, unpublished).

That this technique is also effective in controlling the influence of non-olfactory

cues is seen in the fact that response declines to chance level with progressively increasing dilutions of the test odorant.

The detectability of the n-*aliphatic acetates*

Although few systematic studies of the olfactory stimulating efficiency of acetates have been made, the results, in contrast to those relating to the alcohol and fatty acid series, are in broad agreement. Thus the responsiveness of flies to these compounds (e.g. Cook, 1926; Hughes, 1957) and the magnitude of the slow potentials evoked in the olfactory epithelium of the frog by the first four members of the series (Ottoson, 1958) increase as carbon chain length is expanded. Similarly both Allison and Katz (1919) and Jones (1955) report lower human olfactory thresholds for amyl than for methyl acetate. What appears to be the only conflicting evidence derives from a study by Beck, Kruger and Calabresi (1954) who claim that in man odour intensity in acetates with 6–12 carbon atoms decreases with increasing chain length. However, it is possible that if corrections for vapour pressure were applied to this data the trend would be reversed.

With this exception, then, the present conclusion that the detectability of acetates increases logarithmically as chain length is expanded accords with existing evidence for this series, particularly with that obtained by Hughes on the tsetsefly. There is also a striking similarity between the current results and those of the previous study on alcohols. Thus although the slopes of the lines describing the regressions of concentration at median threshold on chain length differ significantly for the two series ($p < 0{\cdot}001$) the disparity in regression coefficients is nevertheless small (0·09). There is also a significant tendency for the slopes of probit regression lines to decrease in value as both series are ascended.

This conformity suggests that such relations may be of general occurrence in olfaction in the rat, at least in so far as homologous series of aliphatic compounds are concerned. It also confirms the finding that the detectability pattern can be a function of concentration. As was previously suggested this may explain why some workers have found a progressive rise, others an oscillation in odour intensity as a series is ascended.

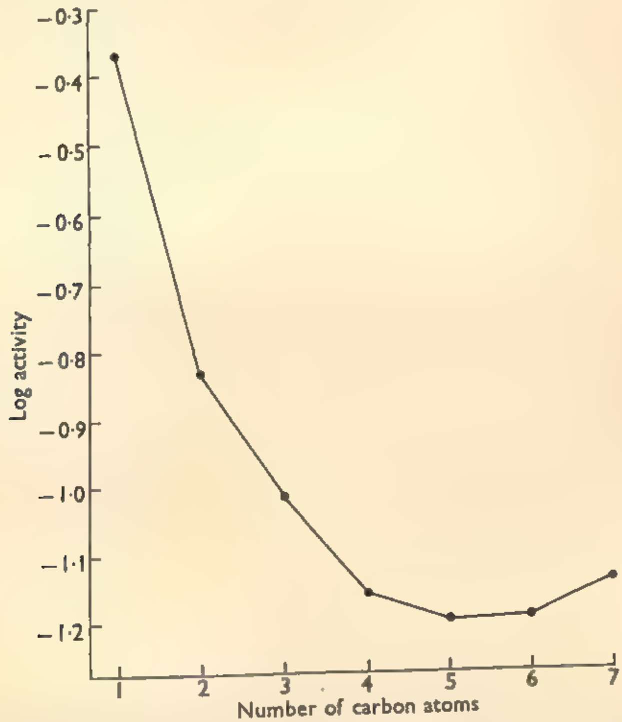
When median threshold concentrations of acetates expressed in molar units (Table I) are compared with the equivalent values for alcohols given in the previous study, the acetates appear to be considerably less stimulating. This is an unexpected result if such properties as solubility and vapour pressure are held to exert any significant influence over detectability. However, this comparison must be viewed with caution since the conditions of the two experiments differed in several respects. In particular, the fact that animals tested on acetates were about 7–9 months older than those used on alcohols may be important. The rat is susceptible to rhinitic infections which generate pathological changes (Kelemen and Sargent, 1946; cf. Clark, 1956), and in one study the volume of the olfactory bulb in male albino rats was found to have decreased by more than 50 per cent. during the second year of life (Smith, 1935). Thus although little is known about the way in which olfactory acuity alters with time, its deterioration in this species at least, seems likely to be marked beyond a certain age. Because of this factor limited value may lie in comparing olfactory thresholds derived from subjects of dissimilar ages.

Analysis of results in terms of thermodynamic activity units

If olfaction in the acetates involves an equilibrium process, median threshold values might be expected to be about equally stimulating when expressed in thermodynamic activity units (see Moulton and Eayrs, 1960). In Figure 4 these values are

plotted against chain length and it is clear that in members containing 4–7 carbon atoms this is in fact the case. However the first four members, although occupying a narrow range of activities, stimulate at decreasing levels as the series is ascended. The pattern is thus closely similar to that derived from the majority of recent studies on olfaction in homologous series, whether in blowflies (Dethier, 1954; Dethier and Yost, 1952), the dog (Moulton, Ashton and Eayrs, 1960), the frog (Ottoson, 1958) or the rat (Moulton and Eayrs, 1960). Whilst the existence of conflicting data relating chiefly to man (see Moulton and Eayrs, 1960) stresses the need for further studies, the weight of evidence now tends to favour the view that in several homologous series olfaction in members of intermediate chain length involves the establishment of an equilibrium, and it can therefore be argued that physical rather than chemical processes control olfaction in these members. Olfaction in the remaining members either does not involve an equilibrium process or it is being obscured by some factor, such as might control the amount of an odorant reaching the olfactory surface.

FIGURE 4



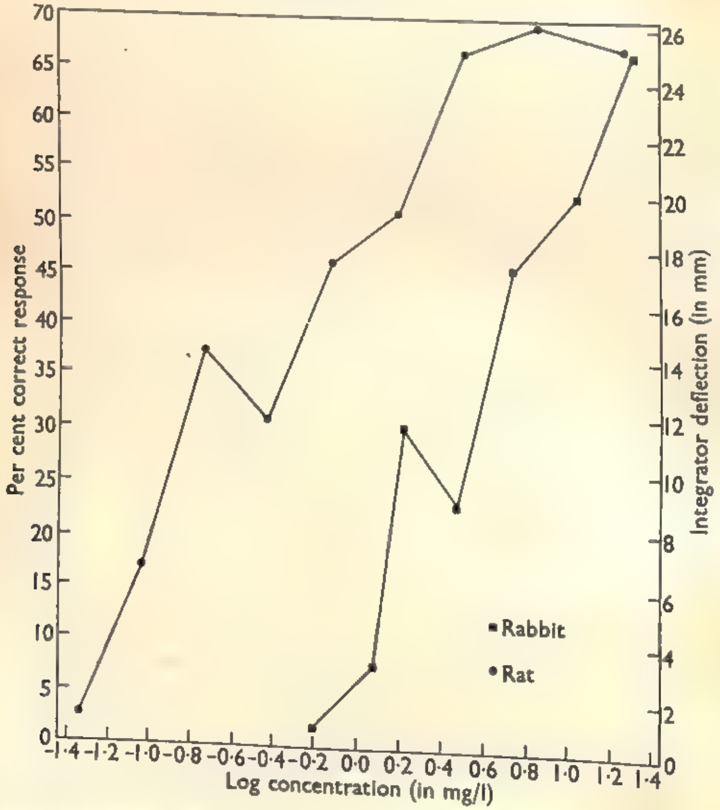
Relation of median threshold concentrations, expressed as thermodynamic activities, to the number of carbon atoms in the alkyl group in an homologous series of acetates.

The stimulus-response curves for acetates

(*a*) *The nature of the anomalies present in the curves.* The most surprising feature of the results is the appearance in the stimulus-response curves of a series of pronounced reversals in slope which bear a linear relation to chain length. Although two or three reversals occur in some curves replications indicate that only one reversal in each acetate above ethyl is related to this trend, and it varies little in appearance throughout the series. Clearly the nature and possible cause of these anomalies are worth considering.

An example of a reversal can be seen in Figure 5 in which the stimulus-response curve for amyl acetate, obtained in the present study, is compared with that derived by Mozell (1958) from recordings of the electrical activity in the olfactory bulb of the rabbit. Although the extreme degree of flattening at the upper margin of the left hand curve, and the low response level at which it occurs, are exceptional, it is, in its general form, representative of data obtained for this series in the present study. Similarly the right hand curve in Figure 5 is considered by Mozell to be typical of his data for amyl acetate, although variations occurred, depending partly on the position of the recording electrode in the bulb. In comparing these two curves it is of interest to note the similarity both in the general shape, and in the position and form of the reversal. This suggests that the anomaly in the present data for amyl acetate, and thus probably in the remaining acetates above propyl which were studied, represent olfactory events.

FIGURE 5



Relation between concentration of vapour of amyl acetate and intensity of response. The behavioural responses of rats, expressed as percentage correct scores, are compared with data for the rabbit derived by Mozell (1958) from recordings of the electrical activity of the olfactory bulb.

However, in the curves for ethyl and propyl acetate, as in those for alcohols containing 5–11 carbon atoms (Moulton and Eayrs, 1960), anomalies occur which are associated with high concentrations of the stimulus. As we have already indicated these are of more doubtful significance since at this level many odorants stimulate non-olfactory receptors and free nerve endings in the nasal mucosa (Parker and

Stabler, 1913; Katz and Talbert, 1930). Such effects may be responsible for Bachem's (1956) claim, which is clearly contradicted by the data of Figure 5, that the sensory curve for olfaction is positively curved throughout its length and exhibits no saturation effect. It may not be so unique among sensory curves as he suggests.

However, the cause of the reversals which conform to the main trend remain to be considered. In the following section a possible explanation will be advanced.

(*b*) *A possible cause of certain anomalies present in the curves.* In recording the electrical activity of the olfactory bulb of the rabbit, Adrian (1953) noted that amyl acetate stimulate both large and small spike units, and he suggested that different groups of receptors were responding. It is thus possible that at low concentrations acetates excite two or more receptor types. As the concentration of a given acetate is increased, however, the firing of one or more receptor types may reach a maximum and then decline, whilst remaining groups continue to show increasing response with increasing concentration. This is consistent with Mullins' (1955) suggestion that some compounds may narcotize certain receptors more easily than others, and could account for a reversal in the stimulus-response curve. However it is also possible that some substances may be bound or adsorbed on to (cf. Beidler, 1954; Moncrieff, 1957), penetrate and accumulate in (cf. Davies, 1953), or otherwise saturate the surface of certain receptors, receptor sites, or phases associated with them, more easily than others. Such effects could lead to a depletion in response but can less readily account for a reversal as opposed to a flattening of the curve.

This may account not only for certain anomalies in the present results and in the curves for several other odorants (Allison and Katz, 1919; Katz and Talbert, 1930) but also for the common observation that the quality of many odours changes on dilution (e.g. Passy, 1892). However, until more evidence is available further speculation on this topic seems premature.

(*c*) *Discrepancies between behavioural and electrophysiological data for amyl acetate.* A further point which stems from a comparison of the curves shown in Figure 5 is that the behavioural data cover a wider range of concentrations than do the electrophysiological data, the discrepancy being about 0·5 log units. Although the value of this comparison is restricted by differences in such factors as the species studied, the number of subjects employed and the use of anaesthetics, a disparity in this direction might nevertheless be expected. The behavioural responses of the intact animal to an odour draw on the firing of a larger pool of olfactory receptors than do those recorded electrophysiologically from one point in the olfactory system, and the probability that receptors of unusual sensitivity are represented will therefore be enhanced. This could account for the fact that the greater part of the discrepancy occurs in the lower segment of the curves.

For these and similar reasons (see, for example, Barlow, FitzHugh, and Kuffler, 1957) it is to be expected that behavioural thresholds should be lower than those obtained electrophysiologically, and it cannot, of course, be implied from Figure 5 that the rat has a lower threshold for amyl acetate than the rabbit. In fact these data suggest that the threshold for the rabbit is about twenty times higher than that of the rat, whereas the equivalent data for the absolute threshold for vision in the cat show an even greater difference. Thus the most sensitive unit in the retina from which Barlow, FitzHugh and Kuffler (1957) recorded, had a threshold which was fifty times higher than that derived behaviourally by Gunter (1951).

In the acetates studied here the asymptotes of the stimulus-response curves are reached in about 2–3 log units of concentration. Mozell's data for a number of odorants show an even smaller range of 1–1·5 log units and he has pointed out that in other senses an increase of 4, 5 or 6 log units of intensity is required (see also

Holway and Pratt, 1936; Bachem, 1956). However, there is considerable variation in this figure for olfaction and for some compounds at least the intensity range may conform more closely to that common in vision and audition (cf. Katz and Talbert, 1930; Moulton, Ashton and Eayrs, 1960).

Conclusions

The present results have confirmed the general findings of the previous report, thus reinforcing the view that the physical properties of a molecule play a dominant role in determining olfactory stimulating efficiency within homologous series and again emphasizing the striking similarity of olfactory responses in the rat and certain species of insects. It is now clear that the rat is a valuable subject for quantitative behavioural studies of olfactory response.

It has also been shown that the stimulus-response curve for olfaction, which has received scant attention in the past, may be more complex in form than is generally assumed and deserves closer study. It may even provide a means of investigating the behaviour of different olfactory receptor types.

My thanks are due to Dr. J. M. Warren for reading the manuscript of this paper and for his assistance in obtaining facilities for the work.

References

Adrian, E. D. (1953). Sensory messages and sensation. The response of the olfactory organ to different smells. *Acta. physiol. scand.*, **29**, 5–14.

Allison, V. C., and Katz, J. H. (1919). An investigation of stenches and odors for industrial purposes. *J. Indust. Eng. Chem.*, **11**, 336–8.

Bachem, A. (1956). The modified Fechner-Weber law as a complex law of dose action. *Amer. J. Psychol.*, **69**, 588–603.

Barlow, H. B., Fitzhugh, R., and Kuffler, S. W. (1957). Dark adaptation, absolute threshold and purkinje shift in single units of the cat's retina. *J. Physiol.*, **137**, 327–37.

Beck, L. H., Kruger, L., and Calabresi, P. (1954). Observations on olfactory intensity. I. Training procedure, methods and data for two aliphatic homologous series. *Ann. N.Y. Acad. Sci.*, **58**, 225–38.

Beidler, L. M. (1954). A theory of taste stimulation. *J. gen. Physiol.*, **38**, 133–9.

Cheeseman, G. H., and Mayne, S. (1953). The influence of adaptation on absolute measurements for olfactory stimuli. *Quart. J. exp. Psychol.*, **5**, 22–30.

Clark, W. E., Le Gros (1956). Observations on the structure and organization of olfactory receptors in the rabbit. *Yale J. biol. Med.*, **29**, 83–95.

Cook, W. C. (1926). The effectiveness of paraffin derivatives in attracting flies. *J. agric. Res.*, **32**, 347–58.

Davies, J. T. (1953). L'odeur et la morphologie des molecules. *Industr. Parfum.*, **8**, 74–9.

Dethier, V. G. (1954). Olfactory responses of blowflies to aliphatic aldehydes. *J. gen. Physiol.*, **37**, 743–51.

Dethier, V. G. (1956). Chemoreceptor Mechanisms. In Grenell, R. G., and Mullins, L. J. (editors). *Molecular structure and functional activity of nerve cells.* Washington, A.I.B.S., 1–33.

Dethier, V. G., and Yost, M. C. (1952). Olfactory stimulation of blowflies by homologous alcohols. *J. gen. Physiol.*, **35**, 823–39.

Eayrs, J. T., and Moulton, D. G. (1960). Studies in olfactory acuity. I. Measurement of olfactory thresholds in the rat. *Quart. J. exp. Psychol.*, **12**, 90–8.

Finney, D. J. (1952). *Probit Analysis.* London.

Gunter, R. (1951). The absolute threshold for vision in the cat. *J. Physiol.*, **114**, 8–15.

Holway, A. H., and Pratt, C. C. (1936). The Weber-ratio for intensitive discrimination. *Psychol. Rev.*, **43**, 322–40.

Hughes, J. C. (1957). Olfactory stimulation of tsetse-flies and blowflies. *Bull. ent. Res.*, **48**, 561–79.

JONES, F. N. (1955). Olfactory absolute thresholds and their implications for the nature of the receptor process. *J. Psychol.*, **40**, 223–7.

KATZ, S. H., and TALBERT, E. J. (1930). Intensities of odors and irritating effects of warning agents for inflammable and poisonous gases. *Technical paper* 480, *U.S. Department of Commerce. Bureau of Mines*, 37 pp.

KELEMEN, G., and SARGENT, F. (1946). Non-experimental pathologic nasal findings in laboratory rats. *Arch. otolaryng., Chicago*, **44**, 24–42.

MONCRIEFF, R. W. (1957). Olfactory adaptation and odor-intensity. *Amer. J. Psychol.*, **70**, 1–20.

MOULTON, D. G., ASHTON, E. H., and EAYRS, J. T. (1960). Studies in olfactory acuity. IV. Relative detectability of *n*-aliphatic acids by the dog. *Anim. Behav.*, **8**, 117–128.

MOULTON, D. G., and EAYRS, J. T. (1960). Studies in olfactory acuity. II. Detectability of *n*-aliphatic alcohols by the rat. *Quart. J. exp. Psychol.*, **12**, 99–109.

MOZELL, M. M. (1958). Electrophysiology of the olfactory bulb. *J. Neurophysiol.*, **21**, 183–96.

MULLINS, L. J. (1955). Olfaction. *Ann. N.Y. Acad. Sci.*, **62**, 247–56.

OTTOSON, D. (1958). Studies in the relationship between olfactory stimulating effectiveness and physico-chemical properties of odorous compounds. *Acta physiol. scand.*, **43**, 167–81.

PARKER, G. H., and STABLER, E. M. (1913). On certain distinctions between taste and smell. *Amer. J. Physiol.*, **32**, 230–40.

PASSY, J. (1892). Sur l'analyse d'une odeur complexe. *C.R. Acad. Sci., Paris*, **115**, 689–90.

SMITH, C. G. (1935). The change in volume of the olfactory bulb of the albino rat during post-natal life. *J. comp. Neurol.*, **61**, 477–508.

BROADBENT'S FILTER THEORY: POSTULATE H AND THE PROBLEM OF SWITCHING TIME

BY

NEVILLE MORAY

From the Department of Psychology, Hull University

Groups of digits were presented binaurally and dichotically to subjects who were asked to recall them. Different rates and patterns of presentation were used, the design being based on that of Broadbent (1954). Broadbent's findings in this field were confirmed. If subjects are presented with simultaneous pairs of digits at a rate of 2 pairs per second dichotically, they cannot recall them if they are asked to alternate between the ears. If however the presentation is staggered, so that although the rate is constant the signals do not overlap, subjects can recall alternately from the two ears. It is suggested that these findings are better interpreted as an interference effect, not a rate effect. Criticism is offered of similar designs to measure "switching rate." Criticism is also offered of Broadbent's estimate of "perception time" in such experiments. An analysis of the quantity and type of errors made by subjects is given, which suggests that Broadbent's theory of a short term store on the peripheral side of a selective filter is in need of revision.

Introduction

Broadbent (1954) presented pairs of digits simultaneously and dichotically to subjects. The signals came in groups of three pairs, and subjects were required to recall the digits either successively (LLLRRR) or by alternating between the ears (LRLRLR). Broadbent found that the latter was only possible providing that the rate of presentation was not greater than one pair every 1½ sec., although if recall were successive the subjects could recall accurately even when the rate was as high as 2 pairs per sec. He interpreted these results to mean that at the faster rates the subject could not switch from ear to ear fast enough, and that the only successful way to deal with the situation was to listen to one ear while the material arriving at the other ear was retained in a short term store on the peripheral side of what he called a "Filter Mechanism."

Now it could be argued that the above experiment is not a fair test of a subject's ability to switch between signals: for what in fact a subject is doing in such a situation is to switch between the signal he hears and a *trace* of the other, not the other signal itself. It seemed that it should give a more genuine measurement of the subject's performance in a switching task if the signals came not in simultaneous pairs, but staggered, so that they never actually overlapped. The following experiment was devised.

Experiment I

Method

Lists of six digits were prepared, and presented to the subjects in groups of ten lists. There is a 10 sec. pause between each group of six during which the subject had to recall. A Brennell Mk. IV two-channel tape recorder was used to present the stimuli to the subjects through headphones, which were Brown moving coil, low impedance type. Subjects were asked to recall the digits either in any order they liked (free recall), or all the digits from one ear followed by all those from the other (successive recall), or by switching from ear to ear (alternate recall). A binaural condition was also used, in which cases the stimuli were obtained by passing the above dichotic lists through a mixing box and then feeding them to both headphones. The free recall conditions were given before any of the other conditions so that all subjects were well practised. The dichotic conditions of presentation were as follows: staggered (‾_ ‾_ ‾_) at 2 signals/ear/sec.; simultaneous (= = =) at 2 signals/ear/sec., and slow overlapping (‾=‾=‾=) at

1½ signals/ear/sec. Half the subjects received the staggered before the alternate stimuli and half vice versa. Sixteen subjects were used, undergraduates and research workers between the ages of 18 and 30.

Results and Discussion

The form of presentation and recall, and the results, are shown in Table I and Figure 1.

TABLE I

Procedure and Results of Dichotic Listening Experiment

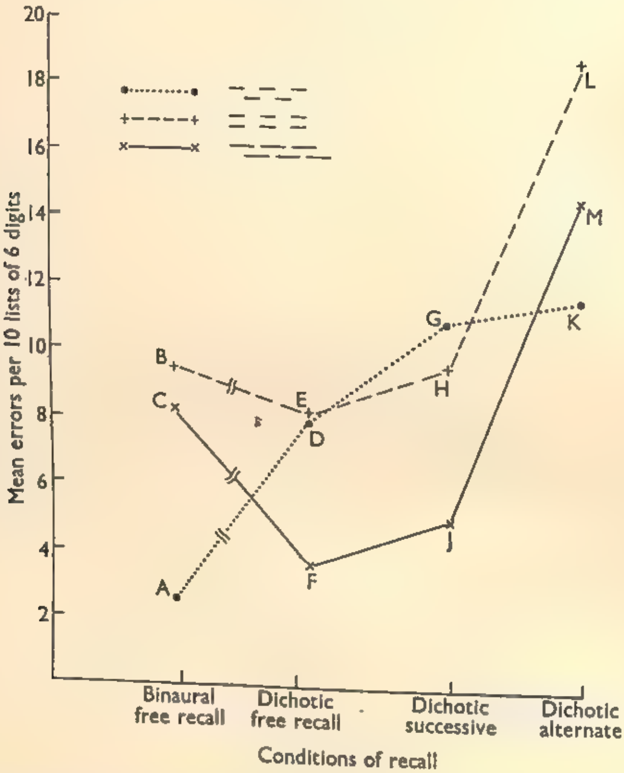
	Presentation	*Recall*	*Rate*	*Mean errors per 10 6-digit lists*	*S.E.*	*Condition*
binaural	successive	free	4/ear/sec.	3·0	0·7	A
binaural	simultaneous	free	4/ear/sec.	9·9	1·3	B
binaural	slow overlapping . .	free	3/ear/sec.	8·0	1·0	C
dichotic	staggered	free	2/ear/sec.	7·6	2·9	D
dichotic	simultaneous	free	2/ear/sec.	7·7	1·7	E
dichotic	slow overlapping . .	free	1½/ear/sec.	3·1	0·8	F
dichotic	staggered	successive	2/ear/sec.	11·3	2·0	G
dichotic	simultaneous	successive	2/ear/sec.	10·0	1·7	H
dichotic	slow overlapping . .	successive	1½/ear/sec.	4·9	1·8	J
dichotic	staggered	alternate	2/ear/sec.	13·5	1·9	K
dichotic	simultaneous	alternate	2/ear/sec.	19·9	2·1	L
dichotic	slow overlapping . .	alternate	1½/ear/sec.	14·6	2·3	M

There is a highly significant difference between successive and alternate recall for simultaneous digits, as Broadbent found: L > H ($p < 0{\cdot}001$ by t test). But there is no significant difference in the case of the signals which were presented in a staggered order in the comparable recall conditions ($p > 0{\cdot}1$). That is, if the signals do not overlap, they can be recalled perfectly alternately even at a rate of 2/ear/sec.—a rate far greater than the limiting rate found by Broadbent for simultaneous digits. It might be thought that staggered signals were inherently easier to perceive than simultaneous ones because of the signals at either end which are "unencumbered" by competing ones; or perhaps that in the simultaneous case the time taken to switch would allow the other half of each pair to fade slightly, but that there is no such advantage is shown by the lack of difference between conditions G and H.

It appears therefore as if the limit on the alternate recall of simultaneous digits is one of interference rather than rate. But rate of presentation does indeed play some part. There is a remarkably constant difference between E, D and F; G, H

and J; and L and M. The only factor common to all these situations is a difference in rate of presentation, and moreover the difference itself is the same in each case. So we may conclude that the difference in performance due to rate of presentation is shown not by the fact that the subjects cannot recall alternately in the simultaneous presentation condition, but by the fact that the $1\frac{1}{2}$/ sec. line is everywhere (in the dichotic situations) lower than the 2/sec. line.

FIGURE 1



Broadbent suggested that the way to estimate switching time was by subtracting the time for two perceptions from the total time for two perceptions plus two switches, the latter being calculated from the subjects' inability to respond to simultaneous stimuli by alternate recall if the stimuli arrived more rapidly than one pair every one and a half sec. He assumed that a perception took about $\frac{1}{2}$ sec. But if we consider condition A in the present experiment, where a binaural condition made by mixing the staggered diochotic stimuli, we have a condition in which the subject was receiving signals at the rate of 4 per sec., and yet the error score is the lowest of the entire experiment. This being so, it is clear that the time taken to perceive a digit must be very much less than 250 millisec., for if it even approached that value we would expect a much higher error score in Condition A. In an unpublished experiment on audioverbal reaction times the writer investigated how much of the signal was needed by the subject to identify a digit, if he knew that the stimulus would be one of the digits "nought" to "nine." An electronic device was used to cut off the signal a predetermined time after it began, and the result showed that there was almost 100 per cent. correct recognition when as little as 30–40 millisec of the signal—whose total length was about 200 millesec. or more—was heard by the

subjects. Now this cannot be used as directly relevant data in the present case, for here the whole of the signal is in fact present, and we do not know how much of it the subject listens to. But it does raise the question of how long a "perception" takes. In the opinion of the writer this question is not one to which it is possible to give an answer. Different sorts of signals may require very different amounts of time to "give rise to a perception," and it is more than likely that had Broadbent used signals of shorter duration than spoken digits he would have obtained data suggesting that the switching time was very much shorter.

Another possible interpretation of the difference between Conditions K and L is this. It may be the case that when signals overlap in time the subject does have to switch from ear to ear; but when they are staggered he listens to both ears simultaneously—as it were "sitting in the middle of his head" instead of "rushing backwards and forwards between his ears." If so, then providing that the signals are staggered, there will never be a decrease in performance with alternating recall unless the signals arrive so rapidly that they would be misheard even in a binaural situation; the suggestion in fact is that even though the stimuli are presented dichotically the subject treats the situation as a binaural one unless they overlap and interfere with one another. In this case the difference between simultaneous and staggered conditions reflects a change in the "strategy" adopted by the brain to deal with different stimulus conditions, but again is not a matter of rate so much as interference. As an alternative to Broadbent's model, for example, one might suggest that when the signals overlap, rather than one set being held up while the other is dealt with, the following happens. The two sets are passed through the filter together but are stored in separate stores according to the source from which they came. During recall they are reproduced store by store.

The above discussion—on the possibility of the subject "sitting in the middle of his head" in a dichotic listening situation—draws attention to another point, one of general interest in any attempt to measure "switching time" by methods such as have recently been used. In none of the experiments so far done on this topic (whether Broadbent's or those of the writer), nor in those which are often invoked in discussions (such as the speech switching experiments of Cherry and Taylor (1954)), is there any guarantee that the subjects *must* have been switching. It is impossible to ensure that a subject is "switching" merely by sending stimuli along what appear to be two separate channels. Whatever our theoretical model may be there is always the possibility that the subject is listening to both ears at once. The following analogy will point the moral. If you are trying to identify people emerging from a house by two doors, there are two strategies at least open to you. Either you can glance rapidly from door to door, in which case the limits on your performance will be due to the rate at which you switch your gaze; or you can steadily fixate a point midway between them, in which case the limiting factor on your accuracy will be the acuity of peripheral vision. But it is impossible to decide which of these things is being done merely by sending people out through the doors at different rates and in different orders. This sort of criticism can be made of all current methods in this field, and it remains to be seen whether ingenuity will triumph over the difficulty or whether we must await a physiological attack on the problem.

We have shown that in some ways an "interference" rather than a "rate" model provides a better account of the limits of performance in these studies. It has also been mentioned in passing that an alternative model without a short term store on the peripheral side of the filter could be designed. Is there any evidence which in fact casts doubt upon this part of Broadbent's model? In his 1954 paper on the perception of simultaneous pairs of digits there is one important score which is not given;

that is the effect of rate on successive recall following simultaneous presentation (LLLRRR following = = =). The following experiment was designed to furnish this data.

Experiment II

Method

Pairs of simultaneous dichotic digits were presented as in Experiment I. Three speeds of presentation were used, 2 pairs/sec., 1 pair/sec., and 1 pair every 2 sec. Subjects were asked to recall successively, that is, all the numbers from one ear followed by all those from the other, starting with whichever ear they pleased. A practice run of ten groups of 3 pairs of digits was given before the experiment started. Half the subjects received the fast condition followed by the medium and the slow speeds, half the reverse. 10 subjects, undergraduates of both sexes, were used.

Results and Discussion

The results are shown in Table II.

TABLE II

The Effect of Rate of Presentation of Simultaneous Digits on Successive Recall

Rate of presentation					
Fast 2 *pairs/sec.*		*Medium* 1 *pair/sec.*		*Slow* 1 *pair/2 sec.*	
Mean	*s.e.*	*Mean*	*s.e.*	*Mean*	*s.e.*
5·55	1·06	4·7	0·89	3·7	1·21

Means are the mean errors per 10 groups of 6 digits for 10 subjects (pooled data).

In his book *Perception and Communication*, Broadbent (1958) stated that the time for which digits could remain in the peripheral store without fading was "of the order of seconds." He has more recently said (Broadbent, 1959), that by this he meant "about a second." Now if this is so, then those digits which have to be stored for more than a second will tend to fade, and we can expect that if the subject must store digits for so long then he will make mistakes during recall. If pairs of digits are presented at the rate of one pair per second, the total time which it takes to present three pairs is rather more than a second, from the beginning of the first to the end of the last (quite apart from any hypothetical "switching time"). So if the subject were letting through, say, the right ear message and storing the left ear message, we would expect the left ear signals to show some signs of deterioration. If we interpret the time constant of "about 1 sec." fairly strictly, we would expect about 1 digit per group of three to show marked signs of deterioration, and moreover, it should be the first one of the "stored" message which shows the greatest deterioration, for this is the one which has been stored the longest. Similarly, if the rate of presentation is greater than this the errors on recall of the "stored" message should be smaller, and those of a slower rate of presentation greater. In the present study the results of the slowest rate of presentation are difficult to interpret, for at a rate of presentation of 1 pair of signals every 2 sec. the subject has time to rehearse what he has already heard. But the data on the other two conditions is clear: the difference, while not being significant, is if anything in the opposite direction to that required by Broadbent's theoretical model. Even if we interpret the time constant of "1 sec." rather

liberally it is strange that the 1 pair per sec. rate of presentation does not show a greater number of errors than the faster rate.

Furthermore, in Table III is given the distribution of errors among the six possible

TABLE III

RELATION BETWEEN ERRORS AND POSITION OF RECALL OF DIGITS IN EXPERIMENT II

Position in recall	*1st*	*2nd*	*3rd*	*4th*	*5th*	*6th*
Speed of presentation						
Fast	5	9	7	12	15	7
Medium..	9	7	4	13	10	4
Slow	2	4	5	8	10	8
F + M + S	16	20	16	33	35	19

Pooled data, 10 subjects.

positions which a digit can occupy during recall. It will be seen that in the overall score it is actually the middle one of the "stored" digits, not the first, which shows the greatest number of errors, although none of these differences (between the 1st, 2nd, and 3rd positions of the "stored" digits) are significant.

Again, the actual total number of errors seems to be too low for the requirements of Broadbent's theory: the average error for each subject in the 1 pair per sec. group works out at only 1·3 digits *per* 10 *groups of* 3 *pairs*, and it would have been lower than this if it had not been for the relatively large scores of one or two subjects. As we said above, it would be expected that the average error for the first of the "stored" digits should be nearer to 10 per group of 10 lists than 1. And in fact if the estimate of "about 1 second" is anything like accurate, *all* the errors are too few.

TABLE IV

ERRORS OF TRANSPOSITION IN RECALL AND ERRORS OF OTHER TYPES IN EXPERIMENT II

Errors of transposition	*Other errors*
44	58

Data from 10 subjects, pooled Fast and Medium conditions.

All the above considerations seem to cast doubt on the existence of a short term memory store on the peripheral side of the filter, in which the input from one ear is held while that from the other is allowed through into the perceptual system. But it does not do more than cast doubt, for the arguments turn on the timing relationships between the signals, and as been pointed out there are two places where there is a great deal of vagueness: firstly, we do not (and probably cannot) know how much of the signals the subject utilizes in identifying which signal he has perceived—that is the "perception time"; and secondly, without knowing more about the exact value of the time constant of the store, we cannot really make predictions about how many digits should be lost in various situations. Indeed, with regard to this latter point, while the inherent vagueness of the postulate remains, it is probably impossible to falsify, although the experiments here described cast severe doubt upon it since neither the number nor the distribution of the errors is what would be predicted.

Finally, in Table IV, there are given the relative number of two different kinds of errors which subjects make during recall. The number of errors which seem to be transpositions between the two ears is compared with the number of errors which seem to be due to mistakes or omissions (that is, if the stimulus was $\begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{matrix}$, mistakes of the kind 423, 156 are compared with, say, 923, 456 or 123, –56). It will be seen that the number of mistakes due to transposition is $\frac{3}{4}$ of the number due to all other causes, and the chance of the transpositions being due to guesses is obviously negligibly small. It is, to say the least, difficult to account for these results on Broadbent's model. If the two inputs are separated before the filter in separate stores then this sort of error could not occur, since the channels are distinct. If the transpositions are supposed to occur after the filter, when the material is all together in the conventional short term memory system, then the number of errors in all is even further reduced to a level now far, far below what would be expected in terms of the time constant given by Broadbent.

In conclusion, it may be said that the present evidence is not strong enough to show that the relevant parts of Broadbent's model are actually wrong; but this is in part at least due to the vagueness of those parts of the model, in particular Postualtes H and J, (Broadbent, 1958, pp. 298, 299). The writer agrees with Broadbent that there is a limit on performance due to switching time, but feels that it has not been measured. He is also a firm admirer of the "Filter Theory" as a general concept. The present paper is more concerned to point out what parts of Mr. Broadbent's lily require gilding than to uproot one of the more attractive blooms in the garden of contemporary psychology.

The writer wishes to thank Professor R. C. Oldfield for providing facilities during the former's time as a research student at the Oxford Institute of Experimental Psychology, during which time half the present work was done. His thanks are also due to the Medical Research Council for a scholarship which enabled him to undertake the work; and to Dr. R. Davis who supervised the research. Lastly he wishes to acknowledge the help and encouragement received from D. E. Broadbent, of the A.P.R.U., Cambridge; and Miss A. Taylor for much stimulating discussion.

References

Broadbent, D. E. (1954). The role of auditory localisation and attention in memory span. *J. exp. Psychol.*, **47**, 191–6.

Broadbent, D. E. (1958). *Perception and Communication.* London.

Broadbent, D. E. (1959). Paper to the Experimental Psychology Society, July, 1959.

Cherry, C., and Taylor, W. K. (1954). Some further experiments upon the recognition of speech, with one and with two ears. *J. acoust. Soc. Amer.*, **26**, 554–9.

AN EXPERIMENTAL SYNTHESIS OF THE ASSOCIATIONIST AND GESTALT ACCOUNTS OF THE PERCEPTION OF SIZE. PART III

BY

R. B. JOYNSON AND N. S. KIRK*

From the Psychology Laboratory, University of Nottingham

Two control experiments are described which are concerned with the effect, on judgements of size, of varying the angular separation of the compared objects in the horizontal plane. The first experiment investigates order effects. The second experiment compares judgements of real size at small and large angles of separation. The findings are discussed in relation to parts I and II of these papers.

Introduction

Parts I and II (Joynson, 1958, *a* and *b*) described an experiment which investigated the effects of angular separation and perceptual attitude on judgements of size. The experiment (afterwards referred to as "the original experiment") was exploratory, and many points require further examination. Two control experiments are described here.

Experiment I

In the original experiment, two main types of judgement were found, "real" (R) and "other than real" (N). Most observers made R judgements first, in order of separation from 40° to 0°; and then N judgements, in order from 0° to 40°. Thus a control experiment is required in which the order is reversed. In the present experiment, N judgements are obtained first in order 40° to 0°, followed by R judgements in order 0° to 40°. So data is obtained complementary to that given in Joynson (1958, *b*, Table I).

A point of special importance which arises is that, in the original experiment, many observers did not seem to be aware of the N possibility initially at 40°. There seemed to be great individual differences in readiness to notice it. Most N judgements at 40° in the original experiment were made only after experience at 0°, where the N possibility is more noticeable. It is therefore especially desirable to determine how far we can obtain N judgements at 40° without this prior experience at 0°.

Apparatus and procedure

The apparatus was identical with that used in the original experiment. Twenty observers, with no knowledge of experimental psychology, compared a 5 cm. rod at 100 cm. with a variable rod at 300 cm. Separations of 40°, 10°, 5° and 0° were again used.

To obtain N judgements, the following instructions were given. They are based on the descriptions of N judgements given in the original experiment. "Do *not* try to make the rods the same real size. They are *not* to be adjusted so that, if measured with a ruler, they would be the same physical length. They are to be made *apparently* equal only. You are *not* to allow for the distance. Judge when they just *look* the same size, *not* when they *are* the same." Each made six judgements at each angle, in the order 40° to 0°, the method of adjustment being used. Observers were instructed to use binocular vision only. They were not told that the angle of separation was to be reduced, nor were they given any information about the possibility of using either simultaneous or successive fixation of the rods (Joynson, 1958, *a*).

* Now in the Department of the Senior Psychologist to the Admiralty.

When they had completed these judgements, they were told: "Now judge when the rods have the same real size, so that if measured with a ruler they would be the same." Each observer made six judgements at each angle, in the order 0° to 40°. They were also instructed here to make a successive comparison, since R judgements in the original experiments were made in this way. At the conclusion of the experiment, observers were again asked to describe what they had been doing as fully as possible.

Initial responses

The first instructions required the observer to make a judgement other than a judgement of real size. Initially, at 40°, there were considerable individual differences among observers in ability to do this. All seemed to consider that a judgement of real size would be fairly straightforward, but not all could see the possibility of something other than this, which would conform to the instructions. A classification is arbitrary, but observers may be roughly grouped as follows. Seven observers seemed to grasp the instructions rapidly and easily (group A: mean = 6·29 cm.). Six understood them, but more slowly and with greater apparent difficulty (group B: mean = 6·19 cm.). Five said they had great difficulty (group C: mean = 5·32 cm.), and it seems doubtful whether they were distinguishing clearly. In these five cases, the difference actually made between R and N judgements was less than 0·5 cm. The remaining two observers (group D: mean = 4·75 cm.) were quite unable to make anything other than a real judgement.

In the case of groups C and D, observers could see the possibility of judging when the rods looked as if they were really the same (i.e. an R judgement), but they could not clearly see any possibility of a judgement contrasted with this, in which the rods just *looked* the same, but were not really equal (i.e. an N judgement). Every effort was made to obtain an other than real judgement, for example by repeating the instructions and explaining them. But this was ineffective. In the case of group C, it seems reasonable to suppose that the small difference which they actually made between R and N judgements, was made, at least in part, because they felt they were expected to make a distinction. Groups A, B and C (eighteen observers) have, however, been combined in the following report, since all made a clear distinction at lesser angles, and all made at least some approximation to an N judgement at 40°. The judgments of these eighteen observers will be described next. The two observers in group D are considered later.

Of these eighteen observers, sixteen made the successive comparison (NSC judgements) at 40° initially, and only two made the simultaneous comparison (NSM judgements). As angle was reduced, an increasing number spontaneously changed from the successive to the simultaneous comparison. Thus at 10° only fifteen made the successive comparison, at 5° only nine, and at 0° only three. As in the original experiment, only approximations to a strictly simultaneous comparison were made at wider angles. When asked to make R judgements, beginning at 0°, a few observers stated initially that the conditions seemed too hard. They were told that they would probably find it easier than they expected, and in fact, encountered no special difficulty.

Quantitative results

The results are given in Table I, and should be compared with Joynson (1958, *b*, Table I and pp. 144–146). Four measures have been calculated: (1) mean values, (2) variance in mean values, (3) mean variability within observers, and (4) variance in mean variability. As in the original experiment, the last three measures show a strong tendency to vary together within each type of judgement, and again they may be taken in combination as a measure of the "reliability" or "consistency" of the judgements. The data have been subjected to the same statistical analysis, in order to assess differences between the three types of judgement at each angle, and to assess trends with angle within each judgement. In addition, the significance of differences between judgements in this experiment and in the original have been calculated (Table II). In the case of N judgements, differences are assessed at each angle, where numbers are large enough to permit comparison. In the case of R judgements, differences are assessed over the four angles in combination. The general pattern of results in the two experiments is very similar. In view of this,

the full statistical results will not be given. The similarities between the two experiments will be stated first, and then the differences.

TABLE I

QUANTITATIVE RESULTS (EXP. I)

Measure (cm.)	*Judgement*	*Angle of separation* 0°	*n*	5°	*n*	10°	*n*	40°	*n*
Mean values	NSM	14·958	15	13·080	9	10·528	3	8·808	2
	NSC	10·534	3	7·537	9	6·921	15	5·683	16
	R	5·648	18	5·392	18	5·240	18	5·118	18
Variance in mean values	NSM	0·168		2·921		1·192		1·109	
	NSC	7·208		0·721		0·944		0·290	
	R	0·953		0·702		0·470		0·368	
Mean variability within observers	NSM	0·174		0·521		0·573		1·037	
	NSC	0·944		0·446		0·346		0·297	
	R	0·374		0·288		0·250		0·252	
Variance in mean variability	NSM	0·0134		0·0695		0·0180		0·3890	
	NSC	0·1210		0·0351		0·0361		0·0233	
	R	0·0180		0·0177		0·0095		0·0088	

(1) *Similarities*

(i) NSM judgements. At 0°, these are not significantly different from 15 cm. ($t = 0{\cdot}40$; $p > 0{\cdot}05$), the correct figure for relative retinal length, and here they show a good level of reliability. They are again strongly affected by angle, and as separation increases they fall below 15 cm. and show a poor level of reliability.

(ii) NSC judgements. These fall between 5 and 15 cm. at 0°, where they again show poor reliability. They are strongly affected by angle, like NSM judgements, though in the opposite way with respect to reliability. As separation increases, they approximate to R judgements in both mean value and reliability, though at 40° they are still significantly greater in mean value than R judgements ($t = 2{\cdot}86$; $p < 0{\cdot}01$).

(iii) R judgements. At 0°, they are significantly greater than 5 cm., the correct figure for equal physical length ($t = 2{\cdot}82$; $p < 0{\cdot}05$), but they are not significantly different from 5 cm. at 5°, 10° and 40° ($t = 1{\cdot}99$, $1{\cdot}49$ and $0{\cdot}83$ respectively, $p > 0{\cdot}05$). They show a good level of reliability throughout, except that variances in mean values are rather greater than in the original experiment. They are comparatively little affected by angle, but, in so far as they are affected, are similar to NSC rather than NSM judgements.

(2) *Differences*

(i) There are marked differences between the two experiments in the proportions of observers making NSM or NSC judgements at wider angles. At 40° in the original experiment eleven of twenty-four made the NSM comparison, as compared with only two of eighteen in this experiment. At 0°, on the other hand, the distribution is very similar: in the original experiment twenty-two of twenty-four made the NSM comparison, compared with fifteen of eighteen in this experiment.

(ii) There is a significant increase in this experiment, in three out of four cases, in variances for NSM judgements at 0° and 5° (Table II).

(iii) There is a significant increase in this experiment for R judgements in both mean value and variance in mean value (Table II).

TABLE II

SIGNIFICANCE OF DIFFERENCES BETWEEN MEAN VALUES IN EXPERIMENT I AND IN THE ORIGINAL EXPERIMENT

Measure (*cm.*)	*Judgement*	*Angle of separation* 0°	5°	10°	40°
Mean values (t)	NSM	0·17 N.S.	1·44 N.S.	—	—
	NSC	—	1·17 N.S.	1·31 N.S.	1.34 N.S.
	R	Overall : 2·71*			
Variance in mean values (F)	NSM	1·20 N.S.	3·01*	—	—
	NSC	—	1·18 N.S.	1·24 N.S.	1·79 N.S.
	R	Overall : 2·38*			
Mean variability within observers (t)	NSM	1·15 N.S.	0·47 N.S.	—	—
	NSC	—	0·65 N.S.	0·24 N.S.	0·66 N.S.
	R	Overall: 0·29 N.S.			
Variance in mean variability (F)	NSM	8·93†	3·51*	—	—
	NSC	—	2·58 N.S.	1·02 N.S.	1·49 N.S.
	R	Overall: 1·22 N.S.			

N.S. = $p > 0{\cdot}05$. * = $p < 0{\cdot}05$. † = $p < 0{\cdot}01$. t_{FB} is used where F is significant.

(3) *Practice effects*

Practice effects for both experiments were assessed by comparing the mean of the observer's first two judgements with the mean of his last two. There was some evidence of moderate trends in the following cases. (i) NSM judgements at 5°: mean practice effect for both experiments combined = + 0·52 cm., i.e. in the direction of reduced constancy (fourteen of twenty observers show the trend). (ii) NSM judgements at 40°: mean practice effect for both experiments combined = + 0·36 cm. (ten of thirteen observers show the trend). (iii) NSC judgements at 0°: mean practice effect for both experiments combined = + 0·30 cm. (ten of fifteen observers show the trend). (iv) R judgements at 0° in the present experiment only (this forms a further difference between the two experiments): mean practice effect = − 0·28 cm., i.e. in the direction of greater constancy (fourteen of eighteen observers show the trend).

It will be seen that practice effects tend to appear where the "reliability" of the judgements is comparatively poor, as might be expected.

Qualitative results

Special emphasis is again placed on observers' reports, and the following should be compared with Joynson (1958, *b*, pp. 146–151).

(i) *N judgements*

There were many comments from observers at 40° initially concerning the difficulty of obeying the instructions and avoiding a judgement of real size. E.g. O.2: "I find it almost impossible to ignore the difference of distance"; O.8: "This is hard work you know.

I feel I have got to fight against taking distance into account"; O.10: "It is very difficult. I have to keep reminding myself not to judge actual size"; O.15: "It is difficult when you can see that the distance is different. I must just try to struggle against it." These comments were not restricted to observers in groups C and D who, as we saw, had particular difficulty in avoiding R judgements here.

As angle decreased, several observers seemed to become increasingly aware of the strength of this tendency. E.g. O.5 (at 10°): "You know, I don't think I'm doing this right. I think I might be half unconsciously doing what you said I was not to do." O.8 (at 10°): "I am just struck by a terrible doubt that I have not been disregarding distance." Finally at 0° they seemed to consider that they were now at last making the judgement with some accuracy, and that earlier judgements were inaccurate, in that they had been unconsciously making much more allowance than they had supposed. E.g. O.2: "I rather suspect that all along I have been making some unconscious allowances for distance, though I was trying not to"; O.8: "I have come to the conclusion my previous judgements were all to hell. I have been progressively more and more conscious of the need to reject the habitual reaction of allowing for distance"; O.11: "I don't know what I was up to at first now"; O.12: "I think, I tended to do that (i.e. allow for the distance) more than I thought at the time." There were also several expressions of surprise at 0°, which again seem to indicate that observers had been allowing for the distance more than they supposed. At 0°, observers again state, as in the original experiment, that the NSM judgement is concerned with "sheer eyesight" or "mere appearance." Here observers are aligning the tops and bottoms of the three images obtained when one rod is continuously fixated.

The two anomalous observers (group D) may be considered here. Both gave judgements approximating to 5 cm. at 40°, 10°, and 5°, and seemed quite unable to make N judgements. At 0°, one suddenly saw the possibility: "Oh! I see I have been compensating for distance. Good god. Fantastic, isn't it? Good lord, amazing." He was, however, still unable to distinguish at wider angles. The second simply made a simultaneous judgement at 0°, approximating to 15 cm., and considered this also was an R judgement. He did not succeed in making a successive comparison at 0°, nor in drawing the distinction between R and N judgements at any angle.

Group C, who had difficulty in distinguishing initially, were asked to repeat N judgements at 40° after experience at 0°. They then gave a mean of 6·82 cm. and now seemed able to distinguish clearly, except for one observer. It will be seen that seven observers of the total twenty (groups C and D: 35 per cent.) could not clearly distinguish the N possibility at 40° prior to experience at 0°, and three observers (15 per cent.) still could not clearly distinguish even after experience at 0°.

(2) *R judgements*

At the conclusion of the experiment, observers were asked to describe the R judgement as accurately as possible. Their descriptions are similar to those obtained earlier. We cannot place emphasis on reports of "allowing for the distance," since this had been mentioned in the instructions. But most observers seemed to find this an appropriate phrase, and there were many comments that the R judgement involves something more than "sheer eyesight," as compared with the N judgement. E.g. O.1: "The apparent judgement doesn't make use of eyesight so much" (i.e. the R judgement makes more use of eyesight).

Most observers again found it difficult to elucidate the nature of this activity, and were intrigued and tantalized by the attempt to do so. O.7: "It is very difficult to say. Language rather fails to describe processes like that"; O.12: "It is difficult to say . . . it is rather queer." Most observers then made comments which suggest that the activity can roughly be called "intuitive looking." O.2: "It is purely instinctive. The distance is taken into account automatically"; O.6: "There just comes a point where I've got to say stop" O.9: "It is just a feeling of rightness." This process involves how the rods "look." O.3: "It is a visual impression more than anything"; O.5: "They just looked right . . . intuition"; O.10: "It is simply when it looked about right." But this involves something more than the "mere appearance" of the N judgement: O.12: "They do not *look* to be identical objects"; O.13: "Something more is involved than just *looking*, but it is subconscious. It is like trying to say how you talk." As this last quotation indicates, the judgement does not seem to involve conscious calculation. As before, comments seem to suggest that the judgement is neither "sheer eyesight," nor a process of conscious calculation, but something in between.

A few observers were again able to go rather further than this, and described the process which was termed "potential looking": that is, they judge when the far rod has that "mere appearance" which they suppose the near rod would have at that distance. O.4: "You look at one, and carrying it in your mind's eye to where the other is, imagine how big it would be there, and then make the other equal to that"; O.16: "I ask myself what would something 'that big' look like at that distance"; O.18: "I have an instinctive idea of how an object the same actual size as the nearer one would appear at a certain distance." Five of twenty observers made such reports, as compared with ten of thirty in the original experiment. Their comments again suggest that in every day life the process would be carried out automatically, and only in the experimental situation do they become reflectively aware of it.

A few observers again mentioned the role of the surroundings. O.15: "In judging the apparent, you forget all about the surroundings, and just look at them as pieces of black. You try to shut everything else out. In the real, you take account of the difference of distance. You allow everything else to influence your judgement. It is much easier." Lastly, in this experiment too, most observers (fifteen of twenty) considered the R judgement to be the more normal and everyday, and again some were most emphatic about this: e.g. O.15: "Very definitely. One doesn't live in two dimensions"; O.16: "Obviously. Nobody every worries about apparent size."

Discussion

Quantitative findings

The quantitative value of *N judgements* is broadly similar, whether they are made in the order 40° to 0°, or the reverse. It might have been expected that N judgements at wider angles would be markedly higher if observers have had prior experience at 0°, than if this experience is absent. But the experiments suggest that any such effect is likely to be of a comparatively minor order. It seems that N judgements deviate from 15 cm. at wider angles because observers cannot help making some automatic allowance for distance. They are able to inhibit this completely if they make the NSM comparison at 0°. But the failure of such experience at 0° to exert a more marked effect at wider angles, emphasizes the automatic character of the allowance. We saw that some of these judgements were tending to lesser constancy with practice. Whether more prolonged practice would be more effective, the experiment does not enable us to state.

Differences between the two experiments do, however, arise in the proportions of observers who make the two types of N judgement. We have seen that many more observers choose to make the NSM judgement at wider angles in the original experiment. The explanation is presumably that, in the original experiment, observers began at 0° by making the NSM comparison, and attempted to continue this as separation increased, feeling it to give more accurate judgements. But in the present experiment, they begin at 40° making the NSC comparison, because the NSM comparison is difficult to notice. It becomes more easily noticed as separation decreases, and observers then change to it. This is again because they feel it to be more accurate, and so at 0° the proportions making the NSM comparison are similar. Thus the effect of different orders of presentation lies primarily in the proportions choosing to make the different kinds of judgement. Its effect on the quantitative value of those comparisons, once they have been made, seems to be of secondary importance. The explanation of the significantly greater variances for NSM judgements at 0° and 5° in the present experiment, is presumably connected with this. It is probably because many observers were here changing from the NSC to the NSM comparison, with a consequent disturbance of their judgements.

The significant increase for *R judgements* in mean value and variance in mean value, as compared with the original experiment, requires explanation. This might be because, in this experiment, R judgements are made immediately after N judgements at 0°. Since observers have been adjusting the variable towards 15 cm. in the

N judgement, they may tend, when making the R judgement, to adjust the variable to a rather greater length than would otherwise seem appropriate. Alternatively, the increase may be a consequence of making R judgements in the order 0° to 40°, rather than in the reverse order.

To decide between these two interpretations, a *supplementary experiment* was conducted with ten additional observers. Using the same apparatus, each made six R judgements at 0° without previous experience of any kind. They gave a mean of 5·20 cm., with variance 0·417 cm. Since this mean is slightly *below* that in the original experiment, and the variance is not significantly different, the first interpretation above is favoured: the greater values in Experiment I are more likely to be a consequence of the immediately preceding N judgements than of making R judgements in the order 0° to 40°. The practice effect in Experiment I for R judgements at 0° is possibly also a function of the preceding experience, since in the supplementary experiment there was no evidence for a practice trend. The further question of why, in both experiments, R judgements are somewhat impaired at 0°, as compared with 40°, is considered in the next experiment.

Qualitative findings

In general, the results substantiate those of the original experiment. The finding that seven observers of the total twenty (35 per cent.) could not clearly distinguish the N possibility at 40° initially, is in good agreement with the conclusion in the original experiment that at least twelve of thirty (40 per cent.) were not aware of it at 40° initially. It seems that, under the circumstances of this experiment, some third of observers find it very difficult, if not impossible, to adopt a so-called "analytic attitude" at wider angles. The reports that, as separation decreased, observers increasingly realized how much they had unconsciously allowed for the distance, and that N judgements are more accurate at 0°, are complementary to those of the original experiment, in which it was said to become increasingly difficult to avoid allowing for the distance as separation increased, and that N judgements were less accurate there.

Experience at 0° seems to facilitate the noticing of the N possibility at 40° in some cases, since of the seven observers just mentioned, four were able to make the N judgement at 40° after experience at 0°. But three (15 per cent.) were still unable to distinguish it. This last finding suggests that in the original experiment, where observers made N judgements at 40° after experience at 0°, a small proportion could probably not in fact distinguish these clearly from R judgements. This suggestion is supported by the observers' comments there, namely, that NSC judgements were not very different from the original R judgements. On re-examining the quantitative results, it has been found that, of the ten observers making the NSC comparison at 40°, three (10 per cent.) differ by less than 0·5 cm. from their R judgement. This is in good agreement with the present experiment.

The observers' descriptions of R judgements confirm those obtained in the original experiment. One earlier conclusion is especially emphasized, namely, that at wider angles observers are less likely to be aware of the activity involved in the R judgement. Seven observers were unable to distingish initially between "when they *are* the same" and "when they just *look* the same." Only as separation was reduced, did they become aware of the activity. This again raises the question whether R judgements differ at different separations.

Experiment II

This experiment is concerned with the comparison of R judgements at 0° and 40°. In both experiments, R judgements are significantly greater than 5 cm. at 0°, and show

greater mean values at 0° than at 40°. They also show poorer "reliability" at 0°, though the trend with angle is significant only in the case of mean variability. This poorer performance in R judgements at 0° is accompanied by several reports of greater difficulty there than at wider angles.

This experiment is concerned with the question whether these are invariable characteristics of R judgements at 0°. If they are, it could be argued (see below) that the situation at 0° is artificial and that the R judgements and reports obtained there are not representative of normal perceptual judgements. It is also possible, however, that R judgements are impaired at 0° because cues to the distance of the farther rod are inferior to those at 40°. At 0°, in both experiments, the stand upon which the farther rod was mounted was almost completely hidden from the observer behind the nearer stand. This might well impair perception of the distance of the farther rod, and reduce constancy. In this experiment, cues to distance are equally good at 0° and 40°, so that we can determine whether in this case R judgements are impaired at 0°. In addition, each observer made thirty judgements, in order to determine whether practice effects occur to any marked extent.

Apparatus

This was the same as in the earlier experiments, except that the far rod now appeared immediately in front of a white cardboard screen, 80 cm. sq. This provided the observer with an excellent indication at 0° of the plane in which the far rod was situated, since it was easily visible beyond the nearer stand. The standard rod was again 5 cm. at 100 cm. distance, and was 0·07 cm. in diameter. The variable was again at 300 cm. and was 0·2 cm. in diameter. The bases of the rods again appeared to the observer in the same horizontal plane.

Procedure

Forty naïve subjects were divided randomly into two groups of twenty. One group made judgements at 0° only; the other group at 40° only. The instructions were as follows: "Judge when the rods are the same real size; that is, state when you think the rods are the same number of inches in length. Please make your judgements carefully, but do not try to calculate in any way. Just rely on your judgement. You will probably find it easier than you might imagine." Observers were also instructed to make a successive comparison of the rods, and to use binocular vision. Each made 30 judgements. The method of adjustment was used, the initial position of the variable being randomly varied between 0·5 cm. and 14·5 cm. No practice judgements were given.

Results

(1) The results are given in Table III. There are small differences between 0° and 40°, in the same direction as in the earlier experiments, but none of these differences is significant.

TABLE III

QUANTITATIVE RESULTS (EXP. II)

Measure (*cm.*)	*Angle of separation* 0°	40°	*Significance*
Mean values	5·028	4·905	$t = 1{\cdot}09.\ p > 0{\cdot}1$
Variance in mean values	0·126	0·116	$F = 1{\cdot}09.\ p > 0{\cdot}05$
Mean variability	0·358	0·331	$t = 0{\cdot}59.\ p > 0{\cdot}05$
Variance in mean variability	0·0264	0·0154	$F = 1{\cdot}71.\ p > 0{\cdot}05$

(2) Since the means are not signicantly different, and since they lie on either side of 5 cm., neither is significantly different from 5 cm., the correct figure for equal physical length.

(3) To assess practice effects, a comparison was first made between the mean of each observer's first and second judgements and the mean of his fifth and sixth judgements. There is no evidence of any important trend. At 0° observers are equally divided in direction of trend shown. The mean trend is − 0·075 cm. (i.e. in the direction of greater constancy). At 40°, ten show a trend in the direction of greater constancy, eight show the reverse trend, and two show no trend. The mean trend is − 0·055 cm.

(4) To assess practice effects over a greater number of judgements, each individual's judgements were next divided into first ten, second ten, and third ten. At 0° the successive means were 5·12, 4·97 and 4·98 cm., and at 40° 4·98, 4·90 and 4·83 cm. Analysis of variance (Table IV) shows a significant overall practice effect at 40° ($F = 5{\cdot}40$; $p < 0{\cdot}01$), but not at 0° ($F = 3{\cdot}07$; $p > 0{\cdot}05$). The analysis also shows (*a*) that there is a significant practice effect for some individuals at 40° ($F = 1{\cdot}82$; $p < 0{\cdot}05$), and at 0° ($F = 3{\cdot}29$; $p < 0{\cdot}001$); and (*b*) that there are significant individual differences at 40° ($F = 18{\cdot}0$; $p < 0{\cdot}001$), and at 0° ($F = 8{\cdot}52$; $p < 0{\cdot}001$).

TABLE IV

ANALYSIS OF VARIANCE (EXP. II)

		0°		40°	
Source	*d.f.*	*S.S.*	*Variance estimation*	*S.S.*	*Variance estimation*
Individuals..	19	74·49	3·92	68·47	3·60
Practice	2	2·82	1·41	2·16	1·08
Interaction	38	17·60	0·46	7·73	0·20
Within individuals ..	540	75·20	0·14	60·91	0·11
Totals	599	170·11		139·27	

(5) Observers' reports did not suggest that there were any important differences between judgements at the two angles.

DISCUSSION

The main finding of the experiment is that, when distance cues are equally good at 0° and 40°, there is little if any difference between R judgements at the two angles. Thus the impairment of R judgements at 0° found in earlier experiments may be ascribed to poorer distance cues there, and is not a necessary characteristic of small angles of separation. There is also no evidence that practice effects are more likely to be found at 0° than at 40°.

Emphasis has been placed on this experiment for the following reasons. We have argued that the R judgements obtained in these experiments correspond closely to our everyday perceptual impressions, and involve an active process. The observers' reports, about "allowing for the distance" and knowing how "mere appearance" varies with distance, have been cited as part of the evidence for this. We have also suggested that the activity is usually carried out automatically, but that observers are more likely to become aware of it at small degrees of separation, where the N possibility is more noticeable (Joynson, 1958, *b*). These conclusions are in contrast with those of the Gestalt psychologists. They hold that our everyday perpetual

impressions involve nothing but a passive "direct experience." They might presumably argue that the observers' reports concerning an activity are primarily derived from 0°; that R judgements here are different in nature from those at 40°; and that the situation at 0° is an artificial one, which throws little light on normal perceptual judgements. Thus Koffka (1935) advocates that the comparison objects should be placed in different directions from the observer, and writes that "if the two objects are too close to each other in the visual field they will influence each other and so distort the picture of uninfluenced constancy." Presumably on this view an impairment at 0° is always to be expected.

It might have been claimed that there was some evidence for this "influence" and "distortion" in the results of the original experiment and Experiment I. But it now seems that the impairment at 0° was a function of our particular experimental set-up, and is not an inevitable occurrence. The results strengthen our conclusion that there is no fundamental difference between R judgements at the two angles, and accordingly the value of the descriptions of R judgements which we obtained does not seem to be impugned. As stated earlier, observers' comments "indicate that they become more aware, at smaller angles, of what they have been doing all the time, not that they change from doing one thing to doing another" (Joynson, 1958, *b*). What happens at 0° is not that the constancy of the everyday perceptual judgement breaks down, but that the possibility of making a different kind of judgement becomes much more noticeable, and the active nature of the everyday judgement becomes evident.

Conclusion

These experiments provide support for the earlier conclusions, and so help to relate the contrasting Associationist and Gestalt accounts of the perception of size. We must in conclusion emphasize a central limitation of the study. The main point of dispute between the rival theories concerns the nature of our everyday perceptual impressions. In these experiments, observers have asserted that R judgements correspond closely to such everyday impressions. But they have also frequently stressed that the experimental judgements are made more deliberately and reflectively than would normally be the case. So caution is necessary in generalizing from the results. It would be desirable to try to obtain less deliberate judgements, in an attempt to approximate more closely to the everyday impression. In conjunction with this, it would be desirable to try to elucidate the conception of real size.

References

Joynson, R. B. (1958, *a* and *b*). An experimental synthesis of the Associationist and Gestalt accounts of the perception of size. (*a*) Part I. *Quart. J. exp. Psychol.*, **10**, 65–76. (*b*) Part II. *Ibid.*, **10**, 142–54.

Koffka, K. (1935). *Principles of Gestalt Psychology*. London.

SHORTER ARTICLES AND NOTES

APPARENT MOVEMENT AND THE BROWN-VOTH EXPERIMENT

BY

JOHN SYLVESTER

From the Institute of Psychiatry, London University

An experiment on apparent movement reported by Brown and Voth and referred to by Osgood and Bartley was repeated. The results obtained differed from those reported by Brown and Voth, but exhibited phenomena not described by them. Further experiments derived from their theories were also carried out, the results failing to support Korte's laws and raising some new points.

INTRODUCTION

In his *Method and Theory in Experimental Psychology* (1953) Osgood gives considerable regard and space to a paper by Brown and Voth (1937) on apparent movement, in which an experiment is described where four lights disposed at the corners of a square are flashed on and off successively, going round the square. Bartley (1941) also devotes some pages to this. According to these authors, when the interval between successive lights in such a set-up is around 150 millisec., a single light is seen to describe a circle. They furthermore claim that this circle has a circumference falling well within the actual square of the flashing lights, as a result of the "cohesive forces" of Gestalt theory. As the interval between lights is further shortened, they say, the circle enlarges again until (at 50 millisec.) four lights are seen simultaneously.

Osgood describes this experiment as "Certainly the most compelling application of this general Gestalt theory of perception" and later says "Careful measurements fully substantiated these predictions from theory as well as many others . . . ." Bartley says much the same. It may be of interest, therefore, to report a repetition of the experiment in this department which has (*a*) failed to obtain the results claimed by Brown and Voth and (*b*) obtained quite different results, namely the occurrence of straight-line Beta movement (optimal movement) in a variety of patterns which accord neither with field theory nor with Korte's laws. Two subjects only were used in the Brown-Voth experiment: fifty-eight subjects were used in the experiments to be described.

Experimentation with the Brown-Voth apparatus was begun with the idea of *using* the Brown-Voth phenomenon (in connection with an experiment on drug effects) and the procedure at first employed differed slightly from theirs; but this, according to their theory, should have made little or no difference to the appearance of the phenomenon. Since, however, the results actually obtained in preliminary trials with the apparatus were quite unexpected, it was decided to reduplicate exactly the Brown-Voth conditions. This introduced difficulties. In the first place, while they give details of light-size and the distances between lights, they do not give the viewing distance (the visual angle, hence, cannot be obtained). There is a further difficulty: they state that "The light boxes A, B, C, D, were duplicated on adjustable grooved arms and with adjustable apertures to allow objective comparison

of the displacements of movement paths with a static field. The light boxes for comparison will be called Ac, Bc, Cc, Dc. The general experimental procedure was to allow S to adjust Ac, Bc and Dc (this was taken to mean that in fact only three comparison lights were used) to phenomenal identity with the certain points of the path of movement as perceived in the field of the main apparatus." Now, if the subject is to reach out and adjust a comparison apparatus adjacent to the display, he cannot be much more than 2 ft. away from it. When the display is a square of side 2 ft. (Brown-Voth experiment III), which give a visual angle of over 53°, the display itself is entirely peripheral (without large eye-movements) and to take in a comparison apparatus as well would be impossible. Large eye movements, of course, destroy the apparent movement under investigation. In view of these considerations, in the experiment to be described the subject was seated 6 or 8 ft. from the display and the comparison apparatus, when used, was adjusted by the experimenter.

A final point: Brown and Voth talk of a "square" but do not say at what angle to the horizontal this square is presented. Their diagrams are drawn with the diagonals of the square horizontal and vertical but their apparatus shows that the angle could be readily adjusted. Bartley (1941) evidently interpreted the presentation as with the square upright. In any case, of course, the plane angle of the square should not affect the results in any way, according to the theory. In the first two experiments described in this paper the presentation was "diamond-form," in the remainder the square was upright.

The first experiment described was the attempt to replicate exactly the Brown-Voth experiment. The others were in fact carried out first.

Method

Experiment 1. Four circular light-boxes (made from 2 in. diameter pill-boxes) were arranged at the corners of a square A B C D of side 6 in. and presented in the form of a "diamond," i.e. with one diagonal vertical. The light-boxes were fitted with 6·3 volt bulbs which could be flashed successively (A, B, C, D, A, etc.) by means of a rotating eccentric arm actuating four contacts equidistantly spaced. The front of each light-box was covered by a lid which had cut in it a circular aperture ¼ in. in diameter. Duration of the lights was controlled by varying the eccentricity of the rotating arm: time interval between lights was varied by adjusting the speed of the motor driving the arm and this motor was geared to a small D.C. permanent-magnet generator which activated a calibrated galvanometer and so continuously indicated the time-interval between successive light stimuli. The timing mechanism was situated in a room other than the experimental one, to prevent the use of auditory cues, and the experiment was carried out in the dark, the experimenter's light being adequately screened. The lights were presented flashing in rotation clockwise with intervals between successive lights (*ti*) of 300 millisec., 250 millisec., 200 millisec., 150 millisec., 100 millisec. and 50 millisec. and the duration of each light (*tl*) was equal to the interval between lights, as in the Brown-Voth experiment. Three subjects were used, a physician, a physicist and an actress. They were seated 8 ft. from the display, and each was given eight trials at each frequency in random order, four with the comparison lights starting on a square greater than 6 in., four with the comparison square starting less than 6 in., alternatively. The experimenter adjusted the comparison lights according to verbal instructions.

Experiments 2–5. Details of these are given with the appropriate experiment in the Results section.

Results

Experiment 1. Table I compares the data given in Table I of Brown and Voth's paper with the data obtained in the present experiment. It will be seen that optimal circularity with a diminution of size was not obtained. What was reported, however, at 100 millisec., was that two lights were moving to and fro along the straight sides of the square, or else that a light was moving up and down the diagonals—optimal, straight-line Beta-movement.

Experiment 2. The comparison apparatus was not used, the experimenter merely saying: "In a moment I shall switch out the room light. I want you to keep looking straight ahead and tell me what you can see—keep up a continuous running commentary, so to speak. Do not try and guess what you *ought* to see but simply describe how things appear to you." The apparatus was then switched on and the speed of the motor operating the contacts slowly increased until well past the subject's report of simultaneity and then slowly decreased again until his report of succession. Five post-graduate psychology students were used as subjects.

TABLE I

SUBJECTS, REPORTS AND ESTIMATIONS OF SIZE OF SQUARE (IN INCHES)

	Time interval between successive lights (in millisec.)					
	300	250	200	150	100	50
Mean	6·28	5·78	5·74	5·41	5·93	5·95
A Prevalent type of movement	Partial on side of square	Optimal on side of square	Optimal with some circularity	Optimal with optimal circularity	Partial simultaneity confused	Simultaneity
Mean ..	6·0	6·0	6·0	5·9	6·1	6·0
B Prevalent type of movement	Succession	Succession	Beta movement jerking round square	Beta round square and D-shaped Beta movement	Beta up and down vertical diagonal, to and fro along horizontal diagonal and to and fro opposite sides of square	Simultaneity

A = Brown-Voth Data. B = Data obtained in this experiment.

Subjects reported seeing, first, four lights appearing and disappearing one after the other successively round a square. Then, as the interval between lights (*ti*) was shortened, they reported that it seemed as though one light was tracing out—rather jerkily—a square. This became smoother as the interval diminished then suddenly two lights in linear Beta-movement were seen, along opposite sides or along the diagonals of the square. Average intervals for the five subjects were (*a*) for one light moving round square (threshold to succession) $ti = 129$ millisec. (*b*) two lights Beta-movement: $ti = 93{\cdot}4$ millisec. (*c*) simultaneity: $ti = 76{\cdot}0$ millisec. Two of the subjects also reported that above the threshold for simultaneity they saw "a sort of whirlpool in a black cloud" (subject 1), "something black swinging round in front of the lights" (subject 2).

Experiment 3. The method was as in experiment 2, except that (*a*) the relative duration of the lights was shortened so that it was half the time interval between two successive lights: (*b*) the light aperture was increased to ½ in.: (*c*) subjects were seated 6 ft. from the display: (*d*) lights were presented at the corners of an upright square. After one or two "up" and "down" runs, the time interval (*ti*) was moved slowly between 20 and 40 millisec., the subjects' reports of the "ghost" being noted. Eight engineering apprentices were used as subjects.

At mean *ti* of 92–64 millisec. (range 113–52 millisec.), all subjects reported Beta-movement, either two lights moving up and down the two vertical sides of the square, or two lights moving to and fro along the horizontal sides of the square, or a light moving alone one or both diagonals.

From time to time the patterning would suddenly change from one form to another in a manner similar to the alternations seen with the Necker cube. Each light appear to move to and fro or "bounce" up and down regularly and smoothly.

At longer intervals, four subjects reported circularity but it was not nearly such optimal movement as the straight-line movement and would be better describable as *phi*. No subject reported a diminution of the circle, though all were specifically asked. Six subjects reported the "ghost"—the black swirling rotation. Some of them said that it appeared not so much "black" as "negative"—a nothingness revolving round the midpoint of the circle.

Experiment 4. The square was increased to a side of 7½ in. and a fixation point was added, in the form of a white light 1 m.m. in diameter at the centre of the square. Also, as it was thought probable that the rotatory after-image phenomenon was to some extent affected by the die-away time of the filament bulbs, four 90 volt neons were substituted for them in this experiment. Ten subjects were used, comprising lecturers, research workers and students in psychology and physiology. The "ghost" phenomenon remained, eight out of the ten subjects reporting it. It seemed to take about two minutes' viewing before its onset. At a *ti* of 40 millisec. and less, the lights appeared quite stationary, not flickering but merely obscured by the passage of a dark-green tinted "sweep" or "one-bladed propeller-like thing" rotating in front of them and sometimes blacking out the fixation light. By reversing the direction of rotation of the lights it was shown that the after-image "ghost" would give adequate cue to the direction of rotation at speeds where the lights themselves had fused.

From 150–250 millisec. of time-interval, subjects reported movement of one light going round a square. At 110–150 millisec. five of them reported one light travelling in a circle but their certainty about this was much less than about the 2-light straight-line Beta-movement seen at shorter intervals. One subject reported that the circle became smaller at $ti = 140$ millisec. This subject saw the circle diminish in size on each occasion of several repetitions but when the comparison apparatus was switched on the effect did not occur—and he saw movement in D-shape, the flat back of the D towards the comparison lights, instead of circularity.

With an interval of 60–110 millisec., straight-line Beta-movement of two lights was seen by nine subjects. All of them experienced the Necker-cube-like alternations between horizontal, vertical and diagonal movement. There seemed to be a tendency for the horizontal movement to occur more readily at intervals nearer to the simultaneity threshold (60 millisec.). Five subjects said, when seeing two lights moving up and down the vertical sides of the square, that the left-hand light seemed to be moving faster than the one on the right, and two of these subjects claimed that this difference in speed was confirmed by the fact that the two lights slowly changed phase, sometimes bouncing up and down together, sometimes alternately. These

two subjects also reported that the right-hand light seemed to be travelling further than the left-hand one.

Experiment 5. The set-up was as in experiment 4 with square of side 6 in., viewing distance 8 ft. and light aperture of 2 in. (i.e. the light-box lids were removed). After one increase-and-decrease, the motor speed was set so that the interval between lights was 100 millisec. and the subject was required to state whenever the phenomenal pattern of the display changed (e.g. when it altered from lights moving up and down to lights moving to and fro). The number of these alternations in a period of 2 minutes was recorded. Subjects were 32 engineering apprentices. Five of these subjects reported a circle getting smaller at an interval time of 100–150 millisec. Two saw no apparent movement of any kind. The mean number of alternations in the two minute period was 13·5. Eleven subjects said that the left-hand light moved faster and /or that the right-hand light moved further. The number of alternations correlated significantly ($r = 0{\cdot}438$) with the E-scale of the Maudsley Personality Inventory, but did not correlate with spiral after-effect persistence time (Holland and Beech, 1958). Twenty-six subjects reported the rotating after-image.

Discussion

There is a footnote to the Brown and Voth paper which says "There were some slight differences between the individual observers in the phenomenal descriptions of the path at the various time-intervals. There also seem to be some day-to-day fluctuations in these. The point to be made is that *it is possible* under certain conditions to *create* the predicted sequences." [My italics.] This point was supported by the experiments described here, since six subjects reported a diminishing circle at around 150 millisec. (though not if the comparison lights were on). It is, however, the only point in their whole paper that these experiments did support. The straight-line two-light optimal movement which nearly all subjects saw seems to be a convincing refutation of their theory of "vector-field-forces." With regard to the seeing of a smaller circle it might be pointed out that the experiments were carried out in the dark, making size-distance estimation difficult except when the comparison apparatus and its adjustment by the experimenter furnished visual and auditory cues. The autokinetic effect certainly operated, many subjects reporting a shift of the whole display in one or other direction.

The results obtained do not accord with Korte's laws, either. Let the square on which the lights are dispersed be called A B C D, and consider it when upright and when two lights are seen moving up and down the vertical sides, the lights actually being flashed successively in a clockwise direction. That the appearance of light C 60–100 millisec. after light B should produce apparent movement from B to C is only to be expected; but light B does not again appear until after an interval five times as long as the B–C interval, when $ti = tl$, or four times as long when $ti = 2\ tl$. Why, therefore, should there be movement from C to B apparently at the same speed as, and quite regular with, the B to C movement? (The movement is strikingly clear and looks like a ball bouncing regularly up and down.) Further, it is light D which comes on after C, yet instead of movement from C to D, at intervals shorter than about 130 millisec. what is seen is movement from A to D. The seen movement from C to B, A to D, etc. is evidently illusory in a different way from normal apparent movement.

The writer is unable to see how these facts can be explained by any sort of cortical field-theory or by satiation-theory (Köhler 1940), although this latter theory is widely held as an explanation of apparent movement. At least as far as the B $\rightarrow$ A and D $\rightarrow$ C or C $\rightarrow$ B and A $\rightarrow$ D movement is concerned, the illusion would seem

to be a mistake in interpretation and hence primarily a function of association-pathways and previous experience rather than of the physiology of the retina or the visual cortex. This type of explanation of apparent movement in general is urged by De Silva (1926), Higginson (1926), Pieron (1952) and Hansel (1953), among others, and is reviewed by Toch and Ittelson (1956).

With regard to the reports of the left-hand light moving faster: there would seem to be a connection here with an aspect of the Kappa phenomenon reported by Cohen, Hansel and Sylvester (1955). Explaining the Kappa effect in terms of previous experience of moving objects they predicted and found that when the light was (apparently) moving downwards it did so as though, to the observer, it was accelerating under gravity.

A similar phenomenon to the rotating negative after-image "ghost" has been reported by Saucer (1954) and called by him "Omega movement." He appears to be unaware of after-image phenomena and, so far as I can make out, puts forward an explanation, instead, which says that movement can be perceived when no object is seen to move! Presumably the reason why the lights appear stationary while the "ghost" is seen to revolve is that during the duration of any light the stimulus from it is accompanied by fused after-images of the other three.

References

Bartley, S. H. (1941). *Vision.* New York.

Brown, J. F., and Voth, A. C. (1937). The path of seen movement as a function of the vector-field *Amer. J. Psychol.*, **49**, 4.

Cohen, J., Hansel, C. E. M., and Sylvester, J. D. (1955). Interdependence in judgements of space, time and movement. *Acta Psychol.*, **11**, No. 3.

De Silva, H. R. (1926). An experimental investigation of the determinants of apparent visual movement. *J. exp. Psychol.*, **37**, 469.

Hansel, C. E. M. (1953). Apparent movement and eye movements. *Brit. J. Psychol.*, **44**, 2.

Higginson, G. (1926). Apparent visual movement and the Gestalt. *J. exp. Psychol.*, **9**, 328.

Holland, H. C., and Beech, H. R. (1958). The spiral after-effect as a test of brain-damage. *J. ment. Sci.*, **104**, 466.

Köhler, W. (1940). *Dynamics in Psychology.* New York.

Osgood, C. E. (1953). *Method and Theory in Experimental Psychology.* New York.

Pieron, H. (1952). *The Sensations.* London.

Saucer, R. T. (1954). Processes of motion perception. *Science*, **120**, 806.

Toch, H. H., and Ittelson, W. H. (1956). The role of past experience in apparent movement. *Brit. J. Psychol.*, **47**, 195.

DISCRIMINATION OF SMALL SHAPES BY THE RAT

BY

P. C. DODWELL*

From the Department of Psychology, Birbeck College

Three experiments on small shape discrimination are described. The first investigates the discrimination of square from circle, a discrimination claimed by Lashley to be extremely difficult for the rat and which has since become a matter of theoretical interest. The second and third test a prediction from the writer's theory of shape recognition; the prediction is refuted.

Introduction

It has been pointed out in a previous paper (Dodwell, 1957) that the interpretation of studies of visual shape recognition in rats (e.g. Fields, 1932; Lashley, 1938) is complicated by part-shape discrimination and by "boundary effects." That is to say, rats tend to discriminate in terms of segments of stimulus shapes, rather than whole shapes, and to discriminate by the relationship of a part of the stimulus shape to the edge of the card on which it is displayed. This may invalidate the use of such studies as confirmatory evidence for theories of shape recognition (Deutsch, 1955; Dodwell, 1957, 1958) although Deutsch and Dodwell both claim that Lashley's results lend support to their (conflicting) theories. An attempt was made to eliminate, or at least reduce, boundary effects in an experiment of the writer's (Dodwell, 1957) and the desirability of experimenting with small shape discriminations was pointed out. The experiments here reported are a first attempt in this direction.

Apparatus and Method

The apparatus used was a Lashley-type jumping stand, modified as previously described (Dodwell, 1957). The subjects were male hooded rats from the colony maintained at the Institute of Experimental Psychology at Oxford, all approximately three months old at the beginning of an experiment. They were put on a feeding schedule one week before training started, being fed for two hours every day, and run after approximately twenty hours without access to food in the first experiment. A slightly different schedule was used for Experiments II and III (see below). Method of handling, recording and scoring were as previously described (Dodwell, 1957). The stimulus shapes used were squares, circles and equilateral triangles, of three different sizes, namely squares of side ¾ in., 1 in. and 1½ in., and circles and triangles with areas equal to those of the three squares. The shapes were outlined in white ink on matt black cardboard, and the rectlinear shapes were always displayed with their lowest side horizontal, unless a different orientation is mentioned. The middle sized shapes were used for training, the others being used to control for discrimination by length of contour, brightness or "subjective size" as previously described (Dodwell, 1957). The rats were trained in the jumping habit on a brightness discrimination. Methodologically, the fact that rats have to be trained to jump before they can be taught shape discrimination is of some importance, since in previous work with the jumping stand pre-training has always been given which could involve discrimination by shape or orientation. Thus, it is usual to train rats to jump initially to an open door, which could involve discrimination of the horizontal base of the rectangular aperture, and hence might predispose the animal to select a shape with horizontal base. In the present experiments the rats were given no training on jumping to open doors, and the brightness discrimination used was such as to preclude any cues of a "shape-like" nature. The rats learned to discriminate a diffuse pattern of white dots on a black card from a plain black card. The dots were not randomly scattered, but concentrated in the centre of the card to form an "aiming point"; from this point the

* On leave of absence at Queen's University, Kingston, Ontario.

gradient of brightness (average separation of dots) was approximately equal in all directions. It might be objected that to pre-train the rats on a brightness discrimination would prejudice them in favour of attempting to discriminate on the basis of brightness, or brightness distribution (Dodwell, 1957) in subsequent training; the first of these possibilities is ruled out by control of brightness in subsequent discriminations, the second cannot be *completely* eliminated, and its possible effects are discussed below. Using the jumping stand technique it is necessary to teach animals to jump by first allowing them to push down the doors, and this involves close contact with the stimulus shapes; in the present experiments shapes are to be presented in such a way that they only subtend a small visual angle at a rat's eyes. It is therefore necessary to have the jumping habit established before proceeding to the discrimination of shapes (from a distance).

Experiments

Experiment I. The purpose of this experiment was to investigate both the rat's capacity for discriminating between small shapes, and Lashley's finding that rats have difficulty in discriminating a square from a circle (a finding shown to be incorrect under certain conditions in the writer's experiment mentioned above). A "small shape" is used here to denote a shape which subtends a visual angle of about 10°, i.e. approximately $\tan^{-1} \frac{1}{6}$ for a shape with diameter 1 in. viewed at about 6 in.

There were ten subjects. On the basis of their performance in the initial brightness discrimination they were divided into two equal groups, group I being trained on the discrimination of triangle from circle, group II on square and circle.

It was found that the discrimination of small shapes under the given conditions requires a large number of trials, and that—within the number of trials given in the experiment—the discrimination is not learned to the criterion of twenty errorless jumps, a criterion previously employed in this type of work by Lashley and others (Lashley, 1938; Dodwell, 1957). Two of the rats (one from each group) had to be discarded because of their poor condition, and one other rat from group II was discarded because of the frequency of badly executed jumps. Training was terminated at the 400th trial for all animals, and their performance over the last fifty trials examined.

A simple χ^2 test on the last fifty jumps for each animal showed that two rats in group I had better than chance scores ($0{\cdot}01 < p < 0{\cdot}05$) and none in group II. It would seem then, *prima facie*, that there is some support for the statement that triangle-circle (group I) is easier than square-circle (group II) discrimination. During training it had seemed that some rats in group II were performing well, so the same test was applied to the previous fifty jumps (Nos. 301–350) for each animal. This showed that two rats in group II had performed at a level "better than chance" (in each case p of χ^2 was between 0·01 and 0·05) although none of the other animals had done so. These results are summarized in Table I, where the individual χ^2 have been

TABLE I

Group	Discrimination		No. of trials	χ^2 Trials 301–350	χ^2 Trials 351–400
	Positive	Negative			
I (4 rats) ..	Triangle	Circle	400	N.S.	$p < 0{\cdot}05$
II (3 rats) ..	Square	Circle	400	$p < 0{\cdot}05$	N.S.

added to give group results. The fact that the performance of rats in group II had deteriorated is of some interest: it may have been due to progressive deterioration in their condition as the experiment progressed (the feeding schedule was somewhat too

rigorous, and was changed in subsequent experiments): however, Wells and Wells (1956) and Sutherland (1957*a*) have also observed falling off of success in a difficult discrimination in the octopus. The level of performance of the animals which achieved better than chance scores on the last fifty trials was not sufficiently consistent to make it likely that the simple transfer tests used in earlier work (Lashley, 1938; Dodwell, 1957) would yield any useful information. It should, however, be possible to estimate transfer effects by giving additional *training* on new discriminations, and measuring relative speed of learning for different pairs of shapes (see experiment II below). In view of the poor condition of the rats, no such further training was attempted.

The results of experiment I may be summed up as follows:—

(1) The discrimination of small shapes is not readily attained by rats in the jumping stand, when boundary effects are reduced and when no pre-training on shape discrimination has been given. When attained, such discriminations may be unstable.

(2) There is no evidence that the rats used could discriminate more easily between triangle and circle than square and circle, although it may be assumed that Deutsch's system of shape recognition, which predicts confusion of square with circle, might be operative under the conditions of the present experiment (*vide* Introduction). The results are compatible with the writer's theory, which makes no firm prediction about the relative difficulty of the two discriminations; either theory might be held to predict difficulty of discrimination of small shapes if there is noise in the "final common cable" (Dodwell, 1957). The experiment supplies evidence incompatible with Deutsch's theory, but gives no specific support to the writer's theory. An attempt to confirm a prediction from the latter theory is made in the next experiment.

Experiment II. A prediction from the writer's theory previously made (Dodwell, 1957) is that a small equilateral triangle (base horizontal) should be confused with an inverted similar triangle, but should be readily discriminable from such a triangle which is positioned with one side vertical. The aim of the experiment is to test this prediction.

There were eight subjects in the experiment. A new schedule was introduced, in which the feeding time was increased to four hours per day; in all other respects the handling and pre-training of the rats were the same as in Experiment I.

All the rats were trained on the discrimination of triangle from circle, but only four showed evidence of discriminating after 400 jumps. These animals had required an average of 393 jumps (range 382–412) to reach a criterion of 18 out of 20 correct, and individual χ^2 over the last fifty jumps had $p < 0{\cdot}01$ for each animal. Thus there can be little doubt that they had fairly mastered the discrimination. The successful rats were divided into two groups on the basis of their performance over the last fifty training trials. Group I was set to discriminate an inverted triangle from a circle; group II to discriminate a triangle with one side vertical (or, what amounts to the same thing, the original triangle rotated through 30°) from a circle. (Simple transfer tests, with both doors unlocked, were not attempted, for reasons indicated above.) It will be recalled that the prediction is that group I will perform significantly better than group II. Each rat was given seventy trials on the transfer discrimination. No rat showed evidence of transferring straight away to either of the transfer triangles, but over the last thirty trials the performance of the two rats in group II was signi-

ficantly above chance expectancy on a χ^2 test ($p < 0{\cdot}05$). These results are summarized in Table II. The results of the transfer training therefore clearly refute the prediction from the writer's theory.

TABLE II

Group	*Discrimination*		*Number successful (average number of trials = 393)*	*Transfers*	χ^2
	Positive	*Negative*			
I (4 rats) ..	Triangle (base horizontal)	Circle	2	Inverted	N.S.
II (4 rats) ..	Triangle (base horizontal)	Circle	2	30°-rotated triangle	$p < 0{\cdot}05$

However, there are several possible explanations of this finding which could obviate the necessity of outright rejection of the theory: (*a*) It could be that in original training on triangles and circles the rats had learnt a "gradient of brightness" discrimination, despite the fact that outline figures were used, rather than "filled in" figures. (The possible confounding effects of using filled in figures in shape discrimination experiments are discussed in a previous paper (Dodwell, 1957).) This would leave unexplained the fact that group II learned the transfer discrimination, since rotation of the triangle through 30° reduces a brightness gradient in the vertical dimension to one very nearly the same as the gradient for the circle. Thus if the rats had been discriminating by means of brightness gradients one would have expected (if anything) less learning in group II than in group I. (*b*) It might be that rotation through 30° is insufficient to disrupt the original discrimination to any great extent. This seems unlikely if the writer's theory holds, and does not accord with other experimental findings on the effects of rotation of shapes (Fields, 1932). (*c*) It might be that the discrimination of the inverted triangle from a circle is for some reason inherently more difficult than the discrimination of a 30° rotated triangle from a circle, so that any facilitation from the learning of the original triangle-circle discrimination would be masked. (*d*) It might be that the shapes used, and the conditions of the experiment (i.e. small shapes against a large homogeneous background) were not sufficient to preclude the complicating factors mentioned in the introduction. If this is the case, it is unfortunate, since rats would certainly have great difficulty in discriminating still smaller shapes.

The third possibility (*c*) seemed to the writer the only one worth investigating further, which is done in experiment III. If this possibility turns out to be wrong, then a definite piece of counter-evidence will have been produced.

Experiment III. There were eight subjects in this experiment. The feeding schedule was the same as for experiment II. The rats were divided into two equal groups at the end of the initial training period, as before. Group I was trained on the discrimination of inverted triangle from circle. Group II was trained on the discrimination of 30°-rotated triangle from circle. The course of learning was similar to that found in experiment II. There were no obvious differences in the

learning rates of the two groups up to 400 trials, when training was terminated, nor did statistical analysis reveal a difference; one can therefore be reasonably confident that the discriminations are of about the same difficulty. This means that the better transfer to the 30°-rotated triangle discrimination in experiment II cannot be explained in terms of the relative difficulty of the two discriminations; hence it must be a transfer effect, which is opposite to the transfer predicted in the writer's theory.

Discussion

The results of experiments II and III cast doubt on the correctness of the writer's theory. No other theory of shape discrimination, so far as the writer is aware, would predict the differential transfer effect of experiment II; it certainly would not be predicted in Deutsch's modified theory (Deutsch, 1958) nor in Sutherland's theory (Sutherland, 1957*b*) even if this latter could be interpreted in terms of contour discrimination.

There is a good deal of evidence to support the contention that a system of shape discrimination similar to that postulated in the writer's theory is operative in the analysis of shapes in simple visual systems (Dodwell, 1957, 1958), which makes it reasonable to look for further evidence before rejecting the theory entirely. The writer has not been able to devise a suitable modification to his theory which would encompass the results here reported.

The writer is indebted to Prof. C. A. Mace for providing experimental facilities, to Dr. J. A. Deutsch for supplying rats and to Mr. N. H. Brown for useful discussions, and for assisting with the running of Experiment II.

References

Deutsch, J. A. (1955). A theory of shape recognition. *Brit. J. Psychol.*, **46**, 30–7.

Deutsch, J. A. (1958). Shape recognition: a reply to Dodwell. *Brit. J. Psychol.*, **49**, 70.

Dodwell, P. C. (1957). Shape recognition in rats. *Brit. J. Psychol.*, **48**, 221–9.

Dodwell, P. C. (1958*a*). Shape recognition: a reply to Deutsch. *Brit. J. Psychol.*, **49**, 158.

Dodwell, P. C. (1958*b*). *Visual Shape Discrimination in Animals and Men.* D.Phil. thesis. Oxford.

Fields, P. E. (1932). Studies in concept formation. I. The development of the concept of triangularity by the white rat. *Comp. Psychol. Monogr.*, **9**, No. 2.

Lashley, K. S. (1938). The mechanism of vision. XV. Preliminary studies of the rat's capacity for detail vision. *J. gen. Psychol.*, **18**, 123–93.

Sutherland, N. S. (1957*a*). Visual discrimination of orientation by octopus. *Brit. J. Psychol.*, **48**, 55–71.

Sutherland, N. S. (1957*b*). Visual discrimination of orientation and shape by the octopus. *Nature*, **179**, 11–13.

Wells, M. J., and Wells, J. (1956). Tactile discrimination and the behaviour of the blind octopus. *Publ. Staz. Zool. Napoli*, **28**, 94–126.

CONTEXTUAL CUES IN SELECTIVE LISTENING

BY

ANNE M. TREISMAN

From the Institute of Experimental Psychology, University of Oxford

Two messages were presented dichotically and subjects were asked to "shadow" whatever they heard on one ear. Somewhere in the middle the two passages were switched to the opposite ears. Subjects occasionally repeated one or two words, at the break, from the wrong ear, but never transferred to it for longer than this. The higher the transition probabilities in the passage the more likely they were to do this. One explanation might be that the "selective filter" (Broadbent, 1958) acts by selectively raising thresholds for signals from the rejected sources rather than acting as an all-or-none barrier.

Introduction

Cherry (1953) found that subjects, when asked to "shadow"—or repeat aloud continuously as they heard it—a passage of prose given to one ear, remained almost completely unaware of the content, though not of the presence of another passage in the other ear; they noticed gross changes of pitch or loudness, but not the introduction of a foreign language or of reversed speech. He concluded that only certain statistical properties of the sounds were analysed, but none of the meaning of the words. Moray (1959) studied in more detail the nature of the attention "barrier." He found that subjects, given a recognition test, failed to show any trace of repeated lists of words given to the rejected ear, and that when they were given a specific set to listen, for example for numbers, they were no more likely to hear them than if they were given more general instructions. The only signals which were sometimes heard were the subjects' own names. He suggested that some kind of analysis was carried out prior to the level of the selective filter, but that only "important" or affective signals were allowed to pass.

It has been shown that dichotic localization of the sound sources is a very effective cue for selective attention (Broadbent, 1954); if both passages are given to both ears it becomes very much more difficult for subjects to separate out the two passages, at least on the first trial (Cherry, 1953). Another factor which plays an important part in "shadowing" tasks is redundancy in the message itself. Moray and Taylor (1958) asked subjects to repeat as they heard them passages composed of statistical approximations to English (made up using Miller's technique (1950)) and found that the number of words omitted was logarithmically related to the order of approximation to English. The higher the transition probabilities between the words, the more likely they were to be heard and correctly repeated.

The aim of the present experiment was to discover whether this type of contextual cue, or expectancy based on transition probabilities between words, would be strong enough to over-ride the dichotic localization cues; whether, if words were made highly probable instead of "important" (as in Moray's experiment), they would also be allowed through the selective attention filter, despite the fact that they came from the rejected ear.

Method

The apparatus used was a Brenell Mark 5 two-channel tape-recorder, the output from each track of the tape going through independent amplifiers to separate earpieces of a pair of headphones. Each subject was asked to equalize the intensities of the two passages.

The passages recorded were fifty words long and were of four different kinds: (*a*) narrative passages from a novel (*Lord Jim* by Conrad), (*b*) extracts from a technical discussion of language (from *Signs, Language and Behaviour* by C. Morris), (*c*) eighth order statistical approximations to English, (*d*) second order approximations to English. There were twelve passages of types (*a*), (*c*) and (*d*) and four of type (*b*). The passages were recorded in pairs, one on each track of the tape, in the following arrangements:—

a–b, b–a, a–c, c–a, a–d, d–a, c–d, d–c, c–c, d–d;

two examples of each pair were used, and the order was randomized. At some point in each recording between the twentieth and the thirty-fifth word, the passages were switched from one track to the other, so that each recording consisted of, on one track, the first part of passage 1 and the second part of passage 2, and, on the other track, the first part of passage 2 and the second part of passage 1. The switches from top to bottom and bottom to top track of the tape-recorder were made to coincide and any pause or change of tone was, as far as possible, avoided. The following example, an "*a–d*" pair, is typical:—

1st track. "While we were talking she would come and go with rapid glances at us
2nd track. "The camera shop and boyhood friend from fish and screamed loudly

leaving on her passage an impression of grace and / is idiotic idea of
singing men and then it was jumping in the tree / charm and a distinct

almost there is cabbage a horse which was not always be the set
suggestion of watchfulness. Her manner presented a curious combination

works every evening is heaviest with bovine eyes looking sideways. . . ."
of shyness and audacity. Every pretty smile was succeeded swiftly by a . . ."

The 18 subjects were undergraduates or research students. They were all given three practice passages of 100 words to shadow. They were then given the passages described above, one track to each ear, and were asked to repeat as they heard it whatever came through one of the two headphones: some were asked to listen to the right and some the left ear.

Six subjects were also given a control series (half of them before the experimental series and half after it). Here the top track was exactly the same as that described above, but on the bottom track a different and irrelevant passage of narrative was recorded, so that the subjects still heard the break in context, but the passage was not switched and continued on the other ear.

Their responses were recorded and later scored for any words repeated from the wrong ear. These were arranged in three categories: intrusions from the wrong ear made within five words before the break in context; intrusions within five words after the break and intrusions elsewhere in the passages. Subjects were also asked if they noticed anything about the rejected passages at the end of the experiment.

Results

There were relatively very few changes to the wrong passage: no subjects changed ears for the whole of the second part to follow contextual rather than localization cues. Three subjects never transferred to the wrong passage. However, fifteen of them did repeat just one or two words on one or more occasions from the rejected ear (the average number of words, for these fifteen subjects, out of twenty 50-word passages being about six). But, when asked afterwards, only one subject had any idea that the passages had been switched to opposite ears, and only two thought they might have said one or two words from the wrong ear. The one who realized the passages had changed sides suddenly commented towards the end of the series, "That time the right ear suddenly wanted to take over." All the other subjects described the rejected passage as "just noise," "perhaps English," or made similar comments. Several noticed that some at least of the passages they were repeating had a break in the context, which they found rather disturbing, and all realized that some (the statistical approximations) were more disconnected than others. Seven subjects were also given, after the other passages, two more where message 1 was in a

woman's voice and message 2 in a man's voice. All of these noticed that the voices switched to opposite ears, although they did not transfer with them in what they were repeating.

The six subjects who did the control series did not insert any of the words which followed after the break in the original passage, so that it is confirmed that in the actual experiment, subjects were genuinely transferring to the other channel and repeating words heard on the side to be rejected.

The number of words repeated from the wrong channel in the "switched" case did seem to bear some relation to the redundancy of the prose. In the results these intrusions are classified by the nature of the material to or from which the transfer was made.

The differences between the conditions were tested, using Student's "t" method, to see which were statistically significant.

Intrusions *from* passage on rejected ear.

After break.	Novel > 8th order.	Significantly different	$p = 0{\cdot}01$	$t = 3{\cdot}2$
	Novel > 2nd order	Significantly different	$p = 0{\cdot}01$	$t = 3{\cdot}6$
	8th order 2nd order	Not significantly different.		
Before break.	None significantly different.			
Not at break.	None significantly different.			

Intrusions *to* passage on accepted ear.

None significantly different, except Before break, 2nd order > 8th order, just significantly different, $p = 0{\cdot}05$, $t = 2{\cdot}6$.

Subjects were significantly more likely to repeat words from the rejected passage after the break if the context they were following was the narrative prose from a novel than if it was a statistical approximation to English, but no difference was seen between the 8th order passages and the less redundant 2nd order ones. There were some, though considerably fewer, intrusions before the break; this was possible because when subjects shadow they are repeating two or three words behind the recorded ones. When all intrusions in the five words preceding the break in context were summed and compared with all intrusions not at the break, divided by eight (to give equal numbers of words which could potentially have been transferred), the difference was statistically significant ($p = 0{\cdot}01$ level). Here, the contextual constraints of the passage which has just come to the right ear from the wrong one seem to work retrospectively, applying transition probabilities in the reverse direction. (Goldman-Eisler (1957) has shown that this is possible, using Shannon's (1949) guessing technique.) There were not enough of these intrusions to show any statistically significant differences between types of prose, although the numbers show the same trend as those coming after the break.

When the results were tabulated in terms of the numbers of intrusions *to* a certain type of passage, there were no significant differences between the different types of prose in intrusions after the break, or intrusions not at the break, perhaps because there were two factors acting in opposite directions: the greater the redundancy the less likely subjects were to leave the correct passage, but also the greater would be the disruption of transition probabilities when the break in context came. In number of intrusions before the break, there were just significantly more from 2nd order approximation to English than from 8th order, presumably an effect of its lower redundancy.

TABLE I

The table shows $\dfrac{\text{Number of intrusions}}{\text{Number of words} \times \text{Number of passages} \times \text{Number of subjects.}}$

1. Intrusions *from* passage on rejected ear.

	(a) *Novel*	(b) *Technical prose*	(c) *8th order*	(d) *2nd order*
After break ..	$\frac{27}{5\times6\times15}$ (6%)	$\frac{10}{5\times2\times15}$ (6·7%)	$\frac{9}{5\times6\times15}$ (2%)	$\frac{8}{5\times6\times15}$ (1·8%)
Before break ..	$\frac{8}{5\times6\times15}$ (1·8%)	$\frac{1}{5\times2\times15}$ (0·7%)	$\frac{2}{5\times6\times15}$ (0·4%)	$\frac{4}{5\times6\times15}$ (0·9%)
Not at break ..	$\frac{3}{40\times6\times15}$ (0·08%)	$\frac{2}{40\times2\times15}$ (0·17%)	$\frac{1}{40\times6\times15}$ (0·03%)	$\frac{8}{40\times6\times15}$ (0·22%)

2. Instrusions *to* passage on accepted ear

	(a) *Novel*	(b) *Technical prose*	(c) *8th order*	(d) *2nd order*
After break ..	$\frac{19}{5\times6\times15}$ (4·2%)	$\frac{9}{5\times2\times15}$ (6%)	$\frac{16}{5\times6\times15}$ (3·6%)	$\frac{10}{5\times6\times15}$ (2%)
Before break ..	$\frac{4}{5\times6\times15}$ (0·9%)	$\frac{2}{5\times2\times15}$ (0·4%)	$\frac{1}{5\times6\times15}$ (0·2%)	$\frac{8}{5\times6\times15}$ (1·8%)
Not at break ..	$\frac{4}{40\times6\times15}$ (0·11%)	$\frac{2}{40\times2\times15}$ (0·17%)	$\frac{2}{40\times6\times15}$ (0·06%)	$\frac{6}{40\times6\times15}$ (0·16%)

A few examples of the kinds of intrusion that occurred were:—

(1) "... I SAW THE GIRL / song was WISHING ..."
... me that bird / JUMPING in the street ...

(2) "... SITTING AT A MAHOGANY / three POSSIBILITIES ..."
... let us look at these / TABLE with her head ...

(3) "... THE GROWL OF THE / "GOAT" (go to) swim fast DURING THE ..."
... book is she went to / thunder INCREASED STEADILY and the ..

(4) "... NEWER techniques will / FOR A MOMENT NOT DARING ..."
... left, while I STOOD / be especially serviceable ...

The words in capital letters were those spoken by the subjects.

Discussion

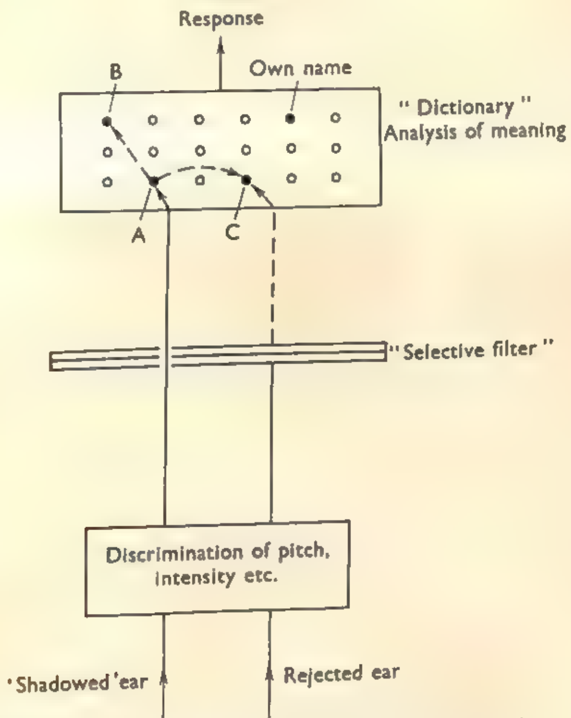
This experiment has confirmed the finding that when the two ears are used as the two channels in a selective listening task, subjects remain almost completely unaware of the content of the rejected passage. Moreover, it seems that contextual cues are not sufficient to make subjects change permanently to the second channel in order to follow the sense of the passage, or even to make them aware of what is being said there. However (using the terms and type of model put forward by Broadbent in his book, *Perception and Communication*) the "selective filter" does seem occasionally to allow one or two highly probable words through from the rejected channel, when the transition probabilities on the correct channel are suddenly contradicted.

The fact that subjects nearly all believe these words came from the same accepted channel makes it seem unlikely that the selective filter has been, as it were, reset momentarily for a different channel. Shadowing experiments suggest that there is a single channel system for analysing meaning, presumably comprising the matching of signals with some kind of "dictionary" and its store of statistical probabilities and transition probabilities gradually learnt through continual use of the language. If this is so, one should be able to avoid the "identification paradox" pointed out by Moray (1959). To explain his finding that subjects sometimes hear their own names when they occur in the rejected message, he suggests that there must be some kind of pattern analysis prior to the filter. Instead, one can suppose that in the "dictionary" or store of known words, some units or groups have permanently lower thresholds for activation, or are permanently more readily available than others: such might be "important" words, a person's own name, or perhaps danger signals (such as "look out" or "fire"); others would be lowered temporarily by incoming signals on some kind of conditional probability basis (along the lines suggested for a learning machine by Uttley (1955)). Thus, for instance, if the three words " I sang a" were heard, the stored trace of the word "song" in the dictionary would have its threshold considerably lowered. The thresholds might of course also be altered in the same way by other types of contextual constraint, by the selection of a verbal category, for example, as in Bruce's experiment (1956) on recognition of words masked by noise. Here the knowledge of the class of words (for example "types of food") allowed them to be recognized at a considerably lower signal-to-noise ratio.

In using the word "threshold" in this context, it is not necessarily meant to imply an intensity threshold, which might be one possibility, but simply that the unit is more or less likely to be activated by incoming signals, or that it is made more or less quickly available. Now if the selective mechanism in attention acts on all words not coming from one particular source by "attenuating" rather than

"blocking" them, that is, it transforms them in such a way that they become less likely to activate dictionary units, it might still allow the above classes of words, with their thresholds which were originally exceptionally low, to be heard. It is suggested that what happens in the experiment described here might be as follows: for the first word after the passages have been switched, two units will be activated in the dictionary, one by the signal from the "selected" ear and the other by the summated effect of the "attenuated" signal from the rejected ear and the lowered threshold due to transition probabilities following the previous word. Either of these units may be chosen for the response, or neither: (many subjects in fact omitted words after the break, perhaps because they had no cues to decide between the two active units). For the second word after the break, the situation is rather different. The transition probabilities will be lowering the thresholds of units following both the units which were last active, but the signals coming in will not be equally effective, since the filter is still operating to favour the selected channel. Thus at either the second or perhaps the third word, the subject will return to the correct ear, and the transition probabilities will be consistent with this until the end of the passage. (Fig. 1.)

FIGURE 1



The thresholds of words B and C are lowered by their high transition probability after word A. Word C is also activated by the "attenuated" signal from the rejected ear and is sometimes heard.

An alternative hypothesis is that the rejected messages are sampled or monitored occasionally. If the signals in the brief sample happened to coincide with a unit in the word-matching system which had been made more sensitive or more available by high transition probabilities, it might emerge in the final response. If not, the competing favoured signal from the selected ear would be the one repeated.

Either of these two possibilities seems a more economical system than any reduplication of analysis before and after the selective filter. Something along these lines also seems necessary to explain why not only a few "important" words, such as one's own name, may be heard from the rejected ear, but also any word which has been made contextually highly probable. If Moray wanted his suggested analysing mechanism prior to the selective barrier to cope with all these possibilities, it would need to be as complex as the one he places after it, at the level of conscious perception.

This might also provide an alternative way of explaining responses to classes of words, such as may occur in perceptual defence experiments. Broadbent (1958) suggests that the filter may be set to select classes of words; but it is difficult to see how a word can be recognized as belonging to a class on the basis of its meaning (the only characteristic common to the class) without already having been analysed individually. An alternative explanation would be that within the "dictionary decoding system" all words belonging to a certain class might have their thresholds raised or lowered relatively to the others. This differs from Broadbent's hypothesis simply in that the selective filter is here confined to acting on "physical" cues of intensity, time or frequency differences, while selection according to characteristics of meaning is done in advance within the analysing or "P" system.

The writer would like to thank Professor R. C. Oldfield who supervized the research, Dr. R. Davis for his helpful criticism and the Medical Research Council for financial assistance. She is also grateful to all the volunteer subjects for their help.

References

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *J. exp. Psychol.*, **47**, 191–6.

Broadbent, D. E. (1958). *Perception and Communication*. London.

Bruce, D. J. (1956). In *Information Theory*. Edited by E. C. Cherry, London.

Cherry, C. (1953). Some experiments on the recognition of speech with one and with two ears. *J. acoust. Soc. Amer.*, **25**, 975–9.

Goldman-Eisler, F. (1957). Speech production and language statistics. *Nature*, **28**, 1497.

Miller, G. A. and Selfridge, J. (1950). Verbal context and the recall of meaningful material. *Amer. J. Psychol.*, **63**, 176–95.

Moray, N. and Taylor, A. M. (1958). The effect of redundancy in shadowing one of two dichotic messages. *Language and Speech*, **1**, 102–9.

Moray, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Quart. J. exp. Psychol.*, **11**, 56–60.

Uttley, A. M. (1955). The conditional probability of signals in the nervous system. *Radar Research Establishment Memo. No.* 1109.

APPARATUS

THE USE OF SMALL FILAMENT LAMPS AS VISUAL STIMULI

BY

A. CARPENTER AND J. A. LEONARD

From the Medical Research Council Applied Psychology Research Unit, Cambridge

The purpose of this note is to draw attention to the fact that the time needed for small filament lamps to reach full brightness is dependent upon the time for which they have been previously switched off. This dependence is a possible source of error in biassed choice reaction sequences, and a circuit is given which avoids the difficulty.

INTRODUCTION

Small low voltage filament lamps are very convenient and cheap sources of visual signals, e.g. in work on reaction times. It is widely realized, however, that there is a delay between the time at which such a lamp is switched on, and its achieving full brightness. In consequence, alternative and perhaps less convenient signals are often used instead. For example, neon lamps have no such delay but are of comparatively low brightness and of a colour which may be unacceptable.

There are cases in which the type of display needed makes the use of small filament lamps almost essential. One such case is the recently introduced "in-line display" in which digits or other characters are projected on to a small screen. This makes use of 6 volt 2 watt lamps with bunched filaments of high surface brightness. In preliminary work on an experiment on choice reaction times, the delay in achieving full brightness with these small lamps has been measured and a circuit devised which ensures that this delay is less than 0·01 sec. in all conditions.

MEASUREMENTS

A lamp from an in-line display unit was mounted close to a photoelectric cell in a dark enclosure, and its output was recorded by photographing a cathode ray tube display onto continuously moving film together with a 100 c.p.s. time marker which was switched synchronously with the lamp. The scale of the final projected image was such that the least time interval which could be noted was 0·001 sec., and the least change in brightness 2 per cent. of the maximum value.

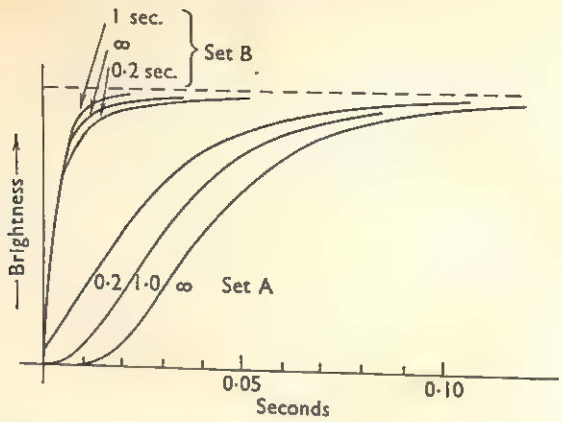
RESULTS

Figure 1 is a superimposition of tracings from the initial measurements, set A, showing the development of brightness with time when the lamp was supplied from a battery of low internal resistance, together with comparable tracings, curve set B, obtained when the circuit of Figure 2 was used. These traced curves are representative of very many obtained, all of which show very close agreement with one another.

The three curves of set A show, first, in the curve marked ∞, the normal switching on of the lamp, after it had been switched off for some ten minutes. It will be seen that the time delay involved is between 0·03 and 0·07 sec., the exact time depending upon how bright the lamp has to be before the subject sees it as "on," in the particular

circumstances of the experiment. This will be the condition in a choice reaction experiment in which the same signal is not repeated directly.

FIGURE 1



If, however, the experiment does involve repetition of the same signal, then the other curves of set A will be applicable. They show that when the lamp has been switched off for only 1·0 sec. or less, then the delay involved in restoring brightness after switching on again is reduced appreciably. When the lamp is switched off for an interval less than 0·2 sec., the corresponding curve does not start from zero. The filament is still hot enough to be emitting light when the circuit is once more made, and it is therefore to be expected that restoration of full brightness would be more rapid. The considerable difference which was found between the three conditions shown in Figure 1 set A was, however, rather unexpected and it shows that there was still appreciable heat in the lamp, even after an interval of 1·0 sec.

In the work which was being planned, it was hoped to use an in-line display in such a manner that individual lamps could be switched on either after an interruption of less than 0·1 sec., or after several seconds, or intermediate cases. Thus it was necessary to devise some means of reducing the resulting variability in the delay.

The circuit shown in Figure 2 was developed, and the curves shown in Figure 1 set B show the behaviour of the lamp in this circuit in the three conditions of set A. It will be seen that 80 per cent. of the full brightness is restored in less than 0·01 sec. in each of the three conditions shown .

The Circuit

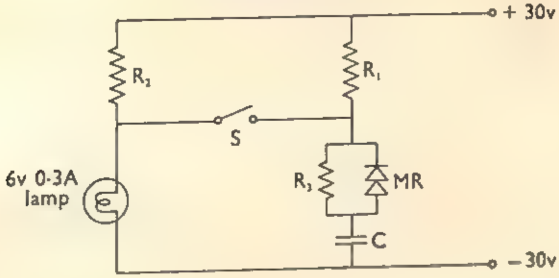
There are three parts to the circuit arrangement shown in Figure 2. Firstly, if R_2, R_3, and MR are excluded, C being joined directly to R_1, a simple condenser arrangement remains. By selection of C to suit a particular value of off period, this simple arrangement is sufficient to ensure a very rapid development of brightness, on switching on. If an experimental arrangement will not demand off periods of less than 1 sec., for example, the other components are unnecessary.

With this arrangement, however, there is still great variability resulting from variation in off period. Short off periods result in appreciable overshoot of brightness, while longer ones result in an initial step up to partial brightness, followed by a much slower approach to the full value.

The second stage in developing the circuit consists in the provision, by R_2, of a small standing bias for the lamp, when the switch is open. This bias of 0·06 amp

develops 0·5 volt across the lamp, and keeps the filament barely visibly red. With this arrangement, the initial rapid cooling of the filament is unchanged, but after about 0·6 sec. there is no further cooling, as is shown by the constant delay on switching on. This circuit would therefore be sufficient if the off period was always longer than 0·6 sec. Shorter off periods result in brightness overshoot.

FIGURE 2



The third development of the circuit is the introduction of MR and R_3, whose effect is to slow the recharge of the condenser C, when S is opened, so that this more nearly matches the decay of brightness of the lamp. It will be seen from Figure 1 that the curve corresponding to an off period of 0·2 sec. no longer shows overshoot, but instead is below the 1·0 sec. curve. These two curves could be made to coincide by the use of a lower value for R_3, but in this case, some overshoot occurs at about 0·5 sec., and the value given is considered to be that more generally useful. The function of MR is to ensure that on closing S and switching on the lamp, R_3 is effectively short circuited.

Conclusion

The circuit of Figure 2 is presented therefore, as an arrangement which will for practical purposes remove the disadvantages of small filament lamps, namely the slow onset of brightness on switching on, and the dependence of this upon the previous off period.

Component values

R_1 .. 100 ohms 7·5 watts.
R_2 .. 500 ohms 2 watts.
R_3 .. 470 ohms 1 watt.
C .. 150 microfarad, 50 volt.
MR .. 30 volt, 1·5 amp. half wave.

Dr. W. E. Hick used a simple condenser circuit in experimental work (Hick 1952) and the present circuit is merely a development of his idea.

Reference

Hick, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11–26.

BOOK REVIEWS

The Annual Review of Psychology, Vol. 11. Edited by P. R. Farnsworth and Q. McNemar. Annual Reviews Inc., Palo Alto, California. 1960. Pp. ix + 544. $7.00.

I used to think that Sidney Smith was joking when he declared "I never read a book before reviewing it; it prejudices me too much." Now that I have finished the *Annual Review* I am not so sure. The principle behind this book is wholly admirable; no one could possibly read for himself more than a tithe of the psychological books and articles which are published every year, and an annual review giving a brief but lucid account of new developments and discoveries in the various branches of psychology would be of immense service to many people—to specialists wishing to keep in touch with departments of the subject other than their own, to students of related fields, and above all to teachers of psychology. But the book itself is very far from admirable; it is constipated with facts, and is rarely relieved by a flash of humour, a felicitous expression or an original thought. Perhaps its shortcomings are as much the fault of the editors as of the contributors. One suspects (rightly or wrongly) that the contributors were told to compress as much material as possible into the small space allowed them, and to keep their own opinions in the background. Comprehensiveness and impartiality are worthy goals, but they are beyond human reach. It is time that the editors bowed to reality, and admitted that the pursuit of forlorn ideals has cost the *Annual Review* a key position which it might well have occupied.

Articles which are too scrupulously impartial are prone to be dull, but dullness is a venial sin. On the other hand, articles which are overly comprehensive are painful to prepare, but leave their readers unenlightened; through trying to say too much, they often fail to say anything properly. Such faults are particularly deplorable in a publication like the *Annual Review* which can only be supposed to interest non-specialists. Specialists could hardly profit from reading articles which summarize two or three years' work in their fields in twenty or so pages. Non-specialists would like to read clear accounts of such recent developments in each branch of the subject as are of general interest. If there are no interesting developments, there is nothing worth reviewing, and if there *are* interesting developments then a review should concentrate on them and describe them fully. Yet the contributors to the *Annual Review*, feeling, doubtless, that they must cover as much ground as possible, have to travel so quickly that they cannot dwell on landmarks, and indeed can scarcely even pause to point out what the landmarks are. They can on the average devote only five or six lines to each publication which they survey; in so small a space it is hardly possible to summarize an article usefully, let alone to add a critical comment which might be helpful to a non-specialist. One article gives eighteen pages of references for twenty pages of text. Several articles contain paragraphs which are simply lists of references, and many contain paragraphs which might just as well *be* lists of references:—

"Optimal numeral form continues to be a subject of interest. Squires (144) was concerned with optimal numeral shape for viewing under red illumination conditions, while Soar (143) compared several numeral shapes under short, 40 msec.-exposure conditions. Each author reached conclusions regarding optimal form as examined under his conditions" (p. 77).

Such paragraphs are quite useless to a non-specialist, who would have neither the time nor the inclination to seek out the articles referred to. A specialist might find the references useful, but then why not simply print the references, suitably classified?

Several authors did not hesitate to introduce obscure technical terms whose meanings they did not have time to elucidate. The sense of this passage may be perfectly clear to an engineering psychologist, but it is not at all clear to me:—

"Chernikoff, Bowen and Birmingham (35) compared tracking performance with a zero-order (positional control) and a fourth-order aided control system when the frequency of the input signal was varied. The results indicated a decided superiority for tracking with the fourth-order aided system for low frequency input signals. However, as course frequency was increased, this superiority was lost, and at high frequencies of input the positional control system became superior. Although these results confound course frequency and system gain for the fourth-order aided control system, they stand as a striking illustration of how quickening can reduce an unmanageable control task to a level which permits highly accurate system performance" (p. 83).

There were only three articles which showed any real signs of breaking out of the general rut, namely, Hoffer's article on "Abnormalities of Behavior," Rotter's on "Psychotherapy," and Pribram's on "Theory in Physiological Psychology." Hoffer has made more effort than most to single out important developments, and his common-sense comments on some of the more pretentious articles reviewed are refreshing. Rotter manages to survey, without breaking into hopeless jargon, even the more esoteric kinds of psychotherapy, which is something of an achievement. Pribram's article is more ambitious than the rest; he attempts to relate recent discoveries in physiological psychology to recent advances in behaviour theory and in the construction of simulated brains. He covers a good deal of interesting material, and is not afraid to speculate. Unfortunately his exposition is in places exceedingly condensed and obscure and his account of simulated brains so short as to be virtually useless. One feels that either he should have been allowed the space to express his ideas more clearly, or else he should have restricted the scope of his article.

If this book accurately reflects the condition and prospects of psychology the outlook is decidedly dismal. It contains no hint of any major advances or revolutionary discoveries; and on the other hand it exemplifies a number of current tendencies which may become dangerous if not speedily checked. Especially worrying is the weight of jargon with which psychologists seem commonly to load their writings. The *Annual Review* contains not merely phrases, or patches, or even paragraphs of jargon, but whole pages of it. Jargon appears to have become the occupational disease of psychologists; its chief symptom is "a flux of words to the pen." The authors of the *Annual Review* are probably no worse stricken than most of the writers whose papers they summarize; but they illustrate a widespread and deplorable tendency towards slipshod writing. Why psychologists in particular should fall victims to the snares of verbiage I do not know; they have no excuse for slovenliness when they can turn for guidance to such masters of exposition as Watson and William James.

Of course every science must have its technical terms, but there is all the difference in the world between introducing a legitimate technical term, and substituting a long word for a perfectly useful short one. What is the point of coining words like "de-emphasizes," "self-devaluator" or "familial" in place of plain English phrases? In the *Annual Review* evidence from the replication of an experiment may be adduced as supportive of the theory which was originally hypothesized, and it may be hoped that as a result of further investigations a more comprehensive systematic integration of the evidence may eventuate. Such words impede the understanding without pleasing the senses.

I am afraid that I may be thought to exaggerate. I shall accordingly illustrate the leading characteristics of jargon by quotations from the *Annual Review*. In each case I have appended a rough translation in brackets; I make no claims that the translations are either perfect or unprejudiced.

1. *The Use of Circumlocutions, Clichés and Polysyllables.*

"The Editorial committee met at the Annual Reviews Building in Palo Alto to attempt long-range planning in the hope that a more adequate topical arrangement of the *Annual Review* chapters might eventuate." (The Editorial Committee met in Palo Alto to work out a new plan for the *Annual Review*.) This is the *very first sentence* of the Editorial Committee's Preface.

2. *Using the Passive Voice instead of the Active.*

"Progress in the scientific prediction of executive success is severely hampered by inability to solve the criterion problem . . . organization-rank and global-effectiveness ratings have served as success criteria in the field studies to date. Organizational rank is questionable on the grounds that non-merit selection considerations contribute unknown, but probably large, amounts of irrelevant variance" (p. 329). (Selection tests fail to predict the success or failure of executives accurately; they fail because most testers take an executive's rank in his organization as an index of his success, and executives who have done badly in tests may yet rise to high rank through having friends on the board.)

3. *Preferring Abstract Words to Concrete Ones.*

"McMurry posits a functional relationship between the structured, routinized character of banking operations and organizations and the personality configurations of most

bank employees" (p. 332). (Personality tests show that bank employees are like banks—careful and methodical.)

It is dangerously easy to wrap the leanness of one's thoughts in a cloak of solemn phrases, and then persuade oneself that such a voluminous garb must hide a substantial body. "Jones and Kohler (65) reported an experiment and its replication that examined the ease with which a subject learns arguments that favor or oppose his position on an issue. They found that the direction of an argument interacted both with the subject's position and the plausibility of the argument. Thus, prosegregation subjects learned plausible prostatements and implausible antistatements better than they learned implausible pro and plausible antistatements. Neutrals learned all types of statements equally well. Jones and Kohler conclude that individuals attempt to protect an attitude position by denying both implausible elements in their own beliefs and plausible elements in opposed beliefs. This conclusion seems plausible" (p. 500). ("We censure others but as they disagree from that humour which we fancy laudable in our selves, and commend others but for that wherein they seem to quadrate and consent with us." SIR THOMAS BROWNE.) "McDavid concluded that these subjects were motivated to conform in order to avoid being different" (p. 486). [*Sic.*]

Not only does jargon solace writers whose thoughts are unremarkable; it is a splendid smokescreen for conceptual confusions. To illustrate: the word "information" has a precise meaning only within the framework of information theory, and outside that framework its meaning is far less exact. Yet these days even when it is used imprecisely it trails clouds of glory which it has captured on its scientific excursions; knowledge of results has become "feedback of correct information" (p. 496), and curiosity behaviour is talked of as a sort of hunger for information (p. 150). It is dangerously easy to pass from the precise to an imprecise use of the word without noticing the transition, and to talk of information as not merely conveyed by stimuli, but as "concentrated" in parts of them (p. 134)—as though bits of information adhered to a stimulus like berries to a bush.

I do not believe that the *Annual Review* faithfully mirrors the state of psychology; it compresses so much into so small a space that the picture it gives is a distorted caricature rather than a true representation. But, as I said at the beginning, a periodical review of progress in psychology *could* perform a valuable service; and the time will soon come when such a review is essential. I would like to offer three suggestions to the Editorial Committee of the *Annual Review*:—

1. They should abandon the ideal of a comprehensive review.
2. They should ask contributors to *select* a limited number of important topics, and to discuss them clearly, critically and at some length.
3. They should ruthlessly eliminate all jargon and all spurious technical terms.

If the *Annual Review* were reformed in these ways it would become a boon to all hard-pressed psychologists, and I should call down blessings upon it.

ALAN GAULD.

Mechanisms of Colour Discrimination. Edited by Y. Galifret. Oxford. Pergamon Press. 1960. Pp. vii + 296. 63s.

This book, the verbatim transcript of an international symposium held in Paris in 1958, is well printed on good paper. Eight of its 296 pages are devoted to a well-ordered, critical and interesting account of the colour vision of insects by K. von Frisch, which is the best presentation of this subject known to me.

Having thus described its serious merits, I can treat the rest of the book with the levity that it deserves.

The participants in the symposium are divided into two classes called, in English, "Reporters" and "Discussants." Only the principal contributions of the former are listed by title in the table of contents, and there is no index. With the exception of an exposition by Hurvich of his opponent-pairs theory, the listed contributions of the Reporters are brief reviews of recent work in fields related, closely or remotely, to colour vision. None but von Frisch's is at the same time orderly, full, critical and in good literary style, and some have not one of these virtues. They are ephemeral lectures, good as lectures, but not designed to be published in a book.

Unmentioned in the table of contents are the minor contributions of the Reporters and those of the Discussants. These vary in length from ten or more pages down to a sentence or single word. Each is printed in the language in which it was uttered, about two-fifths being in English, two-fifths in French and one-fifth in German. Some of the macaronic dialogue provides plesant exercise for the mental agility of the reader. The

onger speeches of the Discusssants either describe unpublished original work, and are in effect hastily-written research papers unprovided with summaries, or are intended to remind the reader of work already published by the speaker. The shorter contributions are mostly genuine discussion, and it is from these that most pleasure and instruction will be gained—instruction not about science but about the scientists. The personalities of the speakers appear vividly in these short speeches. Lively enthusiasm and an unusual quickness of mind is to be seen in almost every paragraph of le Grand, and a subtle sense of humour in all that Galifret says; the characteristic talking-on-paper of Rushton has all three of these qualities. Every reader will be delighted by the exchanges between von Frisch and Medioni on the phototaxis of *Drosophila*, and between Galifret and Medioni on the merits of red, yellow and blue mice as a test for colour vision in cats.

G. S. BRINDLEY.

Le Problème des Névroses Expérimentales. By Jacques Caïn. Bruges. Desclée de Brouwer. 1959. Pp. 168. 150 Fr.b.

The considerable body of knowledge about experimental neurosis accumulated during the last fifty years has been largely the work of independent investigators, who have made little attempt to collate their findings. Pavlov, Liddell, Maier, Masserman, Miller and many others have all studied phenomena which could be labelled as experimental neurosis, but each has used different animals and experimental techniques. In consequence the behaviour shown, the terms used to describe it and the concepts invoked to explain it, have been diverse. In this book Dr. Caïn, who has himself contributed very significantly to the field, draws together this diverse work: to make a synthesis would be impossible, but he presents a clear outline of the problems involved. This he achieves by a careful review of the techniques used (which range from discrimination to audiogenic methods), the types of behaviour shown, the therapeutic measures found useful and the explanatory concepts employed.

In presenting this review he attempts to abstract common features from the varied work he describes. Although he avoids meticulously the pitfalls presented by the overgeneralizations of earlier authors, and the natural temptation to equate behaviour disturbance in animals and man, he does seem to be borne on by a faith that there is a something in common between audiogenic seizures and the consequences of a too-difficult discrimination problem—a faith which the reader may not share. His final conclusion that explanations are not necessarily mutually exclusive, even when they use concepts ranging from "neurophysiological conflict" to "conflicts of meaning," will be well taken, and may also serve to emphasize the breadth of approach necessary for further progress.

R. A. HINDE.

Einführung in die Pharmakopsychologie. By Herbert Lippert. (Vol. 4 of *Enzyklopädie der Psychologie in Einzeldarstellungen.*) Edited by R. Heiss. Bern. Hans Huber. 1959. Pp. 256. Fr.DM 32.

This book represents a useful, but rather unequal, survey of the field which is so very much in the foreground of scientific interest now among both psychiatrists and psychologists. It is written by a German psychologist of no definite school and represents a broad-minded introduction, but not more than that. In some places the author abstracts only the literature; in others there is an attempt at a more critical digestion of results. If one asks what kind of psychology the author represents, one finds that he adheres to a peculiar mixture of German theoretical psychology of the old type, with experimental and biological viewpoints. Hence his chapters on the experimental side are often inadequate, but his reports on clinical topics, such as addiction, are somewhat better. One can hardly blame him that he has little to say in a paragraph on "developmental pharmacopsychology," and the same holds true when he reports on "pharmaco-characterology." However, the list of references to the literature is comprehensive and carefully selected; it contains almost 1,700 items and gives a great deal of work done outside German-speaking countries. It contributes to the value of the book which is certainly to be admired as a courageous attempt to summarize the field from the point of view of the psychologist.

W. MAYER-GROSS.

INDEX OF AUTHORS